Piano Concerto Dataset (PCD): A Multitrack Dataset of Piano Concertos

YIGITCAN ÖZER [iD]

SIMON SCHWÄR [iD]

VLORA ARIFI-MÜLLER [iD]

JEREMY LAWRENCE [iD]

EMRE SEN [iD]

MEINARD MÜLLER [iD]

*Author affiliations can be found in the back matter of this article

ABSTRACT

The piano concerto is a genre of central importance in Western classical music, often consisting of a virtuoso solo part for piano and an orchestral accompaniment. In this article, we introduce the Piano Concerto Dataset (PCD), which comprises a collection of excerpts with separate piano and orchestral tracks from piano concertos ranging from the Baroque to the Post-Romantic era. In particular, using existing backing tracks by the music publisher Music Minus One, we recorded excerpts from 15 different piano concertos played by five interpreters on various instruments under different acoustic conditions. The key challenge of playing along with pre-recorded orchestral accompaniments lies in the exact synchronization of the performer. For guiding the pianists for obtaining a high synchronization accuracy, we used additional click tracks generated with measure and beat annotations of the orchestral tracks, which also are provided in the PCD. The dataset is relevant for a variety of Music Information Retrieval (MIR) tasks, including music source separation, automatic accompaniment, music synchronization, editing, and upmixing.

CORRESPONDING AUTHOR:

Yigitcan Özer

International Audio Laboratories Erlangen, Germany

yigitcan.oezer@audiolabs-erlangen.de

1. INTRODUCTION

Data-driven models for Music Source Separation (MSS) typically require clean, isolated target sources (also called *stems*) for training and evaluation. Research in MSS mainly focuses on separating vocals, bass, and drums from mixtures of popular music songs, mainly due to the availability of multitrack datasets such as MUSDB18 (Rafii et al., 2017) for this task. Separating classical music recordings into individual sound sources has also recently received attention (e.g., Miron et al., 2017; Bugler et al., 2020; Chiu et al., 2020; Gover and Depalle, 2020; Petermann et al., 2020; Chen et al., 2022; Özer and Müller, 2022; Sarkar et al., 2022). Compared to popular music, the constituent sources of classical music recordings often reveal a higher spectro–temporal correlation, which makes the separation task more challenging. Furthermore, the availability of sufficiently large multitrack datasets for Western classical music is a limiting factor for research in this area. In this paper, we introduce a novel multitrack dataset called PCD (Piano Concerto Dataset), which enables both quantitative and subjective evaluation for separating piano concertos.

PCD contains 81 excerpts of multitrack piano and orchestra recordings, each having a duration of 12 seconds. These are selected from 15 different piano concertos from the Baroque to the Post-Romantic period. The variety in the works' complexity, the recordings' acoustical settings, its orchestral instrumentation, and five different performers contributes to the diversity of PCD.

The piano concerto is an essential genre in Western classical music from the Baroque era onward. These compositions are generally written for pianists to demonstrate their virtuosity. Furthermore, the piano concerto has a rich and dynamic sound that is distinctive to this type of music, characterized by opposing musical elements (Cole, 1997). Besides a large number of compositions throughout music history, classical music archives comprise numerous prominent historical, public-domain recordings of piano concertos, which can be useful for various applications in Music Information Retrieval (MIR), including source separation, editing, and upmixing (Martínez-Ramírez et al., 2022), music alignment (Ewert et al., 2009; Prätzlich et al., 2016), automated accompaniment (Dannenberg and Raphael, 2006; Arzt et al., 2008), and audio decomposition (Ewert et al., 2014). We elaborate on the related work for source separation in **Section 2.2**.

To create a multitrack dataset of piano concertos, we use Music Minus One (MMO).[1] MMO provides recordings of backing tracks, in which the lead instrument or the vocal part is omitted, typically the soloist. This allows musicians to practice or perform the solo part along with the pre-recorded accompaniment in case they do not have access to other musicians to play with them. The main difficulty of performing with a pre-recorded accompaniment lies in the absence of any interaction between the player and other musicians. This is particularly problematic for classical music since interpretations can vary greatly in terms of tempo and dynamics. Moreover, piano concertos often contain long sections which only involve orchestral accompaniment. The lack of guidance for the pianist, as typically provided by a conductor, can result in being asynchronous or missing the cue after a long rest. To address this issue, we annotated the measure and beat positions of the backing track of each piano concerto. During the recording sessions, the pianists simultaneously listened to the orchestral accompaniments and sonified click tracks, which were generated using these annotations. Notably, in case of abrupt tempo changes or long piano-solo sections, the additional click tracks have proven helpful for the interpreters while playing along with the pre-recorded accompaniments. Figure 1 displays an excerpt from the Piano Concerto in B Flat Minor, Op. 23, 1st Movement by Peter Ilyich Tchaikovsky and its recording process. The recording sessions are followed by post-production for generating cohesive mixtures of newly recorded piano tracks and existing MMO accompaniments. As a main contribution of PCD, we provide dry and reverberant recordings of piano and orchestra stems and their mixtures.

The remainder of the article is organized as follows. In **Section 2**, we give an overview of existing multitrack datasets and investigate relevant MSS applications. In **Section 3**, we address the role and significance of piano concertos in Western classical music and review their form and compositional structure. As the main contribution of this article, we describe the content of PCD in **Section 4** and outline its recording process and challenges. In **Section 5**, we describe the different interfaces for accessing the dataset. In **Section 6**, we provide an exemplary usage of PCD for separating piano concertos with a baseline U-Net model presented by Özer and Müller (2022). Finally, we conclude in **Section 7** with prospects on the potential applications of the PCD.

2. RELATED DATASETS AND THEIR APPLICATION IN MUSIC SOURCE SEPARATION

2.1 DATASETS

For the training and evaluation of data-driven models, datasets constitute an essential component of MIR research. In particular, the availability of multitrack datasets has led to impressive results of data-driven MSS approaches that focus on separating popular music recordings. In the Western classical music domain, several datasets have been introduced for polyphonic vocal music (McLeod et al., 2017; Schramm and Benetos, 2017; Cuesta et al., 2018; Rosenzweig et al., 2020),
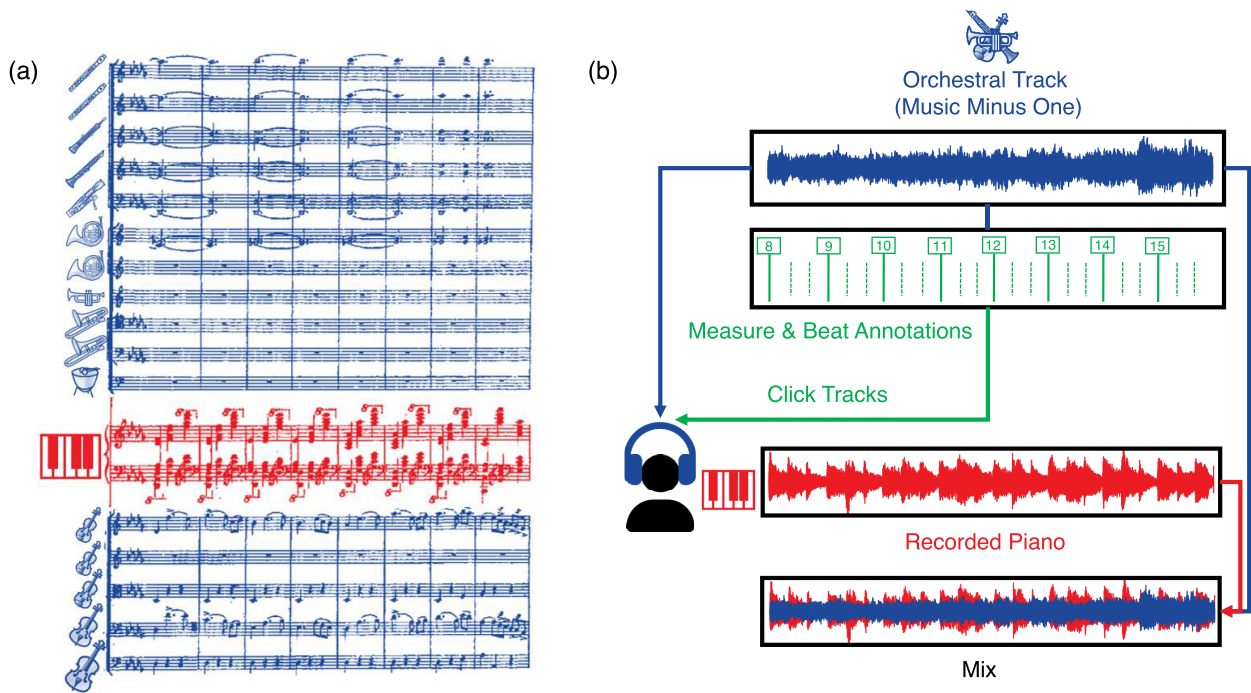
Figure 1 Overview of the recording process. **(a)** The sheet music of measures 8–15 from Tchaikovsky's Piano Concerto No. 1 in B Flat Minor, Op. 23, 1st movement. **(b)** During the recording process, the pianist is supposed to play synchronously with the backing track. In a real-life recording process, the pianist and conductor interact for optimal synchronization and cohesion between the piano and orchestra. In our scenario, however, playing along with a pre-recorded accompaniment is a difficult task for the performer. To address this challenge, they listen to metronome-like click tracks sonified using measure (solid green) and beat (dashed green) annotations in addition to the orchestral accompaniments. As the result of a final mastering step, PCD comprises dry and reverberant *synchronous* recordings of piano and orchestra accompaniments selected from 15 different piano concertos and their mixes.

NAME & AUTHOR	#R	DUR
WWQ (Bay et al., 2009)	1	00:09:00
TRIOS (Porter, 2012)	5	00:03:12
PHENICX (Schedl et al., 2016)	4	00:10:36
Bach10 (Miron and Martorell, 2017)	10	00:05:30
URMP (Li et al., 2019)	44	01:36:00
EnsembleSet (Sarkar et al., 2022)	9	01:03:34
PCD	81	00:16:12

Table 1 Multitrack instrumental datasets in the Western classical music domain, indicating the number of recordings (#R) and their total duration (Dur) in hh:mm:ss.

which comprise isolated recordings of vocal ensembles. For instrumental music, however, there are only a few multitrack datasets (see Table 1). Bay et al. (2009) presented the Woodwind Quintet (WWQ) dataset, which includes separate tracks of a woodwind adaptation of Beethoven's String Quartet, Op.18 No. 5. The TRIOS dataset (Porter, 2012) involves multitrack recordings of four classical pieces and one jazz piece, as well as their transcriptions. The PHENICX-Anechoic dataset (Schedl et al., 2016) comprises annotations and audio material of anechoic multitrack recordings of four orchestral works, which differ in terms of the number of instruments per instrument class. Bach10 (Miron and Martorell, 2017)

consists of multitrack recordings of ten chamber music pieces where each work comprises four parts (SATB) played by violin, clarinet, saxophone, and bassoon. Li et al. (2019) introduced the University of Rochester Multimodal Performance (URMP) dataset, which addresses the music performance as a multi-modal art form and provides the musical score, as well as the audio recordings of the individual stems of 44 ensemble pieces. Their work also describes the challenges of maintaining synchronization and musical expressiveness while creating a multitrack dataset of classical music pieces. Sarkar et al. (2022) presented the EnsembleSet, which consists of synthesized multitrack recordings of strings, woodwind instruments, and brass, generated by using MIDI files from RWC (Goto et al., 2002) and Mutopia.[2] For an overview of a variety of publicly available datasets in MIR, we refer to Bittner et al. (2019).[3]

2.2 MUSIC SOURCE SEPARATION (MSS)

MSS is defined as the task of separating a recording of multiple instruments or voices into individual musical sound sources. Generally, a musical source may refer to singing, an instrument, or an entire group of instruments, such as an ensemble or orchestra. Isolating individual musical sources contained in a sound mixture is useful in a variety of applications, including creating karaoke systems, assisting in music production, enabling music transcription, and supporting music analysis. Due to the

non-stationary spectro-temporal properties of music signals and the high correlation of constituent sound signals in a music recording, MSS is a challenging task (Cano et al., 2019). In the last years, deep neural networks (DNNs) have led to major improvements in separating musical sources (e.g., Jansson et al., 2017; Takahashi and Mitsufuji, 2017; Stoller et al., 2018; Stöter et al., 2019; Hennequin et al., 2020; Défossez, 2021; Kim et al., 2021). A main pre-requisite for supervised MSS models is data availability, as training DNNs requires large datasets with clean, isolated recordings.

Unlike in popular music production, where individual instruments are often recorded separately, the direct interaction between musicians is typically an essential aspect of the recording process for classical music. When musicians perform together in the same room, they have more flexibility in adjusting tempo and dynamics, resulting in a more cohesive and expressive performance. As a result, there are hardly any multitrack recordings available for classical music.

To circumvent the problem of missing multitrack training samples, artificial training data has been created by random mixing as a data augmentation method. For example, Chiu et al. (2020) generated training material by mixing classical violin and pop piano solo recordings for the separation of piano and violin duos. For the quantitative evaluation of the MSS model, they then used 16 multitrack piano and violin recordings from MedleyDB (Bittner et al., 2014). Özer and Müller (2022). also generated artificial training material by randomly mixing sections selected from the solo piano repertoire (e.g., piano sonatas, mazurkas) and orchestral pieces without piano (e.g., symphonies) to train an MSS model for separating piano concertos in a lead-accompaniment separation setting. In this scenario, the lack of multitrack recordings made the quantitative evaluation of the MSS model difficult. PCD will enable the subjective and quantitative evaluation of MSS models addressing the separation of piano concertos, providing a wide range of works recorded by various performers in different acoustic environments.

3. PIANO CONCERTOS IN WESTERN CLASSICAL MUSIC

As it is a central theme in PCD, we highlight in this section the compositional structure and evolution of piano concerto as a genre of central importance in Western classical music. A piano concerto is a musical composition written for piano and orchestra. It typically consists of multiple movements, with the piano playing the primary role and the orchestra providing the accompaniment. Since the Baroque period, piano concertos have been composed by many composers from all epochs until today. As a result, piano concertos are an enduring

and popular form of classical music and continue to be enjoyed by audiences around the world.

In the seventeenth century, the earliest use of the term *concerto* in Western classical music referred to its literal meaning *combined effort*. The "combined effort" sense persisted until Johann Sebastian Bach, whose keyboard concertos depend on the *reconciliation* of cembalo or harpsichord and other instruments (Cole, 1997). One has to consider that in J. S. Bach's time, the keyboard did not yet have the status of a virtuoso instrument as it does today. When it appeared in association with other instruments, it was initially associated with the term *continuous bass instrument* (Schering, 1905). Nowadays, pianists often perform baroque keyboard concertos on the modern piano.

Whereas the high Baroque period cultivated various kinds of concertos, the *solo concerto*, which comprises a lead instrument accompanied by an orchestra, emerged as the preeminent type of this form in the high Classical period. In the late eighteenth century, the classical concerto evolved to an independent form, incorporating form-functional elements associated with the Baroque period, e.g., the ritornello, and the Classical period, e.g., the classical sonata form (Caplin, 1998). The pioneers of the Vienna Classic, Haydn, Mozart, and Beethoven, wrote piano concertos that involve a dialogue between orchestra and solo instrument (Schering, 1905).

During the course of the nineteenth century, romanticism brought a new interest in orchestral color, and composers explored a variety of sounds obtained by closely intertwining the solo instrument and the orchestra. Additionally, the piano had grown in tonal capabilities compared to its usage in the Baroque and Classical periods. As a result, Romantic piano concertos diverged from the Classical form (Girdlestone, 1948). For example, the focus of the interaction between orchestra and piano shifted in favor of the soloists in the case of piano concertos by Frédéric Chopin. In contrast, while renouncing the mere virtuoso display of the soloist, Robert Schumann's Piano Concerto in A minor, Op. 54, is considered a masterpiece of thematic and melodic integration of piano and orchestra. Romanticism reached a climax in Brahms' piano concertos, interchangeably splitting the themes between orchestra and the piano. Finally, the virtuoso style in the Romantic period witnessed its best examples in Tchaikovsky's famous first piano concerto (Op. 23), but even more by the post-romanticism in the piano concertos by Rachmaninov.

4. PIANO CONCERTO DATASET (PCD)

In this section, we present the PCD as our main contribution of this article. In **Section 4.1**, we cover the details of the musical content and characteristics of the dataset. We define the naming conventions of the

included files in **Section 4.2**. In **Section 4.3**, we explain our approach for the alignment of pianists with the pre-recorded orchestral tracks. We elaborate on the recording process in **Section 4.4**, describe the required pre-processing steps in **Section 4.5**, and finally the post-production in **Section 4.6**.

4.1 DATASET CONTENT AND CHARACTERISTICS

This section describes several aspects concerning the content and characteristics of PCD. The dataset consists of 81 excerpts selected from 15 piano concertos by 10 different composers, as shown in Table 2. Here, the *WorkID* specifies the prefix of each filename in the dataset encoding the composer, assigned work number (i.e., Op, BWV, and KV), and the movement, respectively. For further information on the naming conventions of the audio and annotation files in the dataset, see **Section 4.2**. In addition to various compositional styles ranging from the Baroque to the Post-Romantic era, PCD includes different difficulty levels of piano concertos. For example, J. S. Bach's Piano Concerto in F Minor, BWV 1056, is classified as moderately difficult, whereas Rachmaninov's Piano Concerto No. 3 in D Minor, Op. 30, is a very challenging virtuoso work for pianists.

Although we recorded longer sections, including the exposition, development, or sometimes entire movements of piano concertos, we decided to extract and provide only shorter excerpts of the recordings for several reasons. First, practicing and performing entire movements can be difficult for pianists. Second, it is time-consuming for sound engineers to edit and process longer recordings. Third, depending on the compositional style, piano concertos may involve long sections where the piano and orchestra do not play together, which does not serve the multitrack dataset. We will make the raw piano recordings (also of longer sections) available, upon request.

The choice of excerpts is mainly based on musical coherence and a balance between piano and orchestra. Besides passages where both parts play together, we also included sections where the piano and orchestra follow a conversational style, such as in Beethoven's Piano Concerto No. 4 in G Major, Op. 58. In order to account for a suitable duration of the excerpts, we regarded two guidelines. First, the excerpts need to be long enough to involve a complete musical phrase. Second, they should be relatively short for their usability in a subjective listening test. Based on these criteria, we

WORKID	COMPOSER	FULL NAME	MVM	PID	#V	#E	DUR
Bach_BWV1056-01	J. S. Bach	Piano Concerto in F Minor, BWV 1056	1	YO	2	10	120
Beethoven_Op015-01	Beethoven	Piano Concerto No. 1 in C major, Op.15	1	MM	1	6	72
Beethoven_Op019-01	Beethoven	Piano Concerto No. 2 in B Flat Major, Op. 19	1	ES	2	4	48
Beethoven_Op037-01	Beethoven	Piano Concerto No. 3 in C Minor, Op. 37	1	ES	2	4	48
Beethoven_Op037-02	Beethoven	Piano Concerto No. 3 in C Minor, Op. 37	2	LR	1	1	12
Beethoven_Op058-02	Beethoven	Piano Concerto No. 4 in G Major, Op. 58	2	ES	2	2	36
Chopin_Op021-03	Chopin	Piano Concerto No. 2 in F Minor, Op. 21	3	ES	1	5	60
Grieg_Op016-01	Grieg	Piano Concerto in A Minor, Op. 16	1	ES	1	1	12
Mendelssohn_Op025-01	Mendelssohn	Piano Concerto No. 1 in G Minor, Op. 25	1	ES	2	2	24
Mozart_KV414-01	Mozart	Piano Concerto No. 12 in A Major, KV.414	1	YO	1	2	24
Mozart_KV467-01	Mozart	Piano Concerto No. 21 in C Major, KV.467	1	YO	1	5	60
Mozart_KV467-02	Mozart	Piano Concerto No. 21 in C Major, KV.467	2	YO	2	6	72
Rachmaninoff_Op018-01	Rachmaninov	Piano Concerto No. 2 in C Minor, Op.18	1	JL	1	5	60
Rachmaninoff_Op018-02	Rachmaninov	Piano Concerto No. 2 in C Minor, Op. 18	2	JL	1	5	60
Rachmaninoff_Op018-03	Rachmaninov	Piano Concerto No. 2 in C Minor, Op. 18	3	JL	1	5	60
Rachmaninoff_Op030-01	Rachmaninov	Piano Concerto No. 3 in D Minor, Op. 30	1	ES	2	6	72
Saint_Op022-01	Saint-Saëns	Piano Concerto No. 2 in G Minor, Op. 22	1	ES	1	2	24
Schumann_Op054-01	Schumann	Piano Concerto in A Minor, Op. 54	1	ES	2	4	48
Tschaikovsky_Op023-01	Tchaikovsky	Piano Concerto No. 1 in B Flat Minor, Op. 23	1	ES	2	6	72
				Σ		81	972

Table 2 Overview of the dataset indicating the work identifier (WorkID), composer, full name of the work, movement (Mvm), performer identifier (PID), number of versions (#V), number of excerpts (#E), and total duration in seconds (Dur). The versions here refer to distinct performances recorded under different acoustic conditions and played on different pianos.

decided on a duration of 12 seconds. This audio length has been a good compromise for musicality while being the longest recommended duration for Multiple Stimulus with Hidden Reference and Anchors (MUSHRA) listening tests (Series, 2014).

For a wide range of interpretations, five pianists participated in the curation of the dataset: Emre Şen (ES), Jeremy Lawrence (JL), Lisa Rosendahl (LR), Meinard Müller (MM), and Yigitcan Özer (YO). All the performers have provided their consent to publish the recorded material for research purposes under a Creative Commons license. The performers' skills range from amateurs (LR, MM) to semi-professional players (JL, YO) to a concert pianist (ES), and their experiences differ accordingly. LR is a historian and musicologist, and MM is a full-time professor in MIR, playing the piano as a hobby. Among the semi-professional performers, JL is a Master's student in electrical engineering with a strong musical background and experience as a pianist, whereas YO is a Ph.D. candidate working on MIR, with a background in electrical engineering and piano performance. ES is a concert pianist who regularly performs recitals and plays piano concertos with orchestras.

Furthermore, the room acoustics vary among the performances, ranging from a small and relatively dry domestic space (R2), via small recital halls (R1 and R3), to a spacious concert hall environment (R4). Each room is also associated with a different grand piano model. Table 3 summarizes the differences in recording conditions for each room.

In addition to distinct acoustic conditions, PCD includes recordings that vary in quality and orchestral accompaniments. The recordings of Rachmaninov's Piano Concerto No. 2 in C Minor, Op. 18, performed by JL are considered the highest quality recordings in the dataset. These performances were recorded in multiple sessions and underwent exhaustive post-processing. Moreover, this is the only instance where the orchestral accompaniment is synthetic (as provided by MMO), whereas other backing tracks are real recordings. Note that we also provide recordings of multiple movements from the same piece for three works: Beethoven's Piano Concerto No. 3 in C Minor, Op. 37, Mozart's Piano Concerto in C Major, KV.467, and Rachmaninov's Piano Concerto No. 2 in C Minor, Op. 18. Furthermore, there are two versions of certain excerpts, providing different piano recordings using the same underlying orchestral accompaniments.

To gain a more comprehensive understanding of the statistics of the dataset, the distribution of the number of pieces per composer is presented in Figure 2a. In PCD, Rachmaninov is the most prominent composer, with 21 excerpts and a total duration of 252 seconds. Beethoven comes in second place, with 17 excerpts, followed by Mozart, J. S. Bach, and Tchaikovsky. Figure 2b provides an overview of the number of excerpts played by each performer. Most of the performances are by the concert pianist ES. Note that several pieces were performed by the same performer in different rooms. For example, ES performed Tchaikovsky's Piano Concerto No. 1 in B Flat Minor, Op. 23 both in R3 and R4. Figure 2c illustrates the number of excerpts per room. The majority (32) of the recordings took place in R4, roughly a quarter of them (21) in R3, 15 in R1, and 15 in R2.

ROOM ID	ROOM DESCRIPTION	PIANO	#E	DUR
R1	Lecture Hall (Fraunhofer IIS)	Yamaha C3	15	180
R2	Private Studio (Jeremy Lawrence)	Yamaha C3X	15	180
R3	Music Academy (Emre Şen)	Seiler	21	252
R4	Saygun Concert Hall (Bilkent University)	Steinway D	30	360
		Σ	81	972

Table 3 Overview of different rooms where the recordings took place, number of excerpts recorded in each room (#E), and their total duration in seconds (Dur). Note that the piano model is different in each acoustic environment.
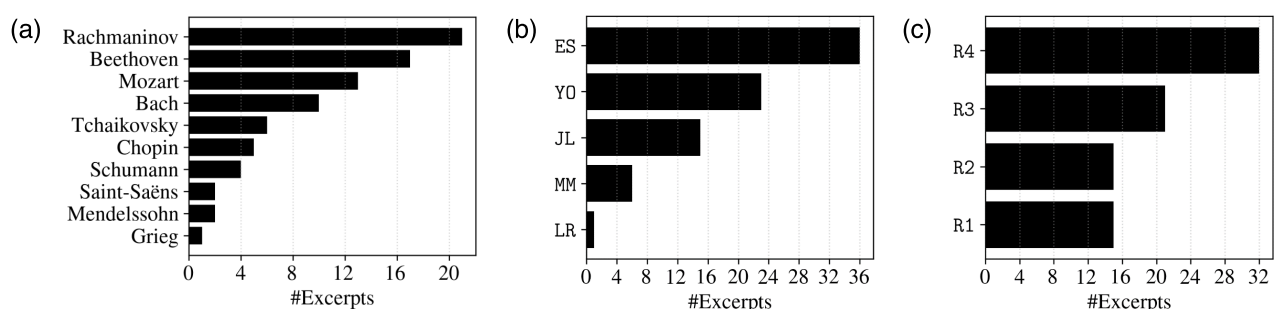

Figure 2 Various bar plots describing the dataset. The number of selected 12-second excerpts is indicated by the horizontal axis per **(a)** composer, **(b)** performer, and **(c)** acoustic environment.

4.2 NAMING CONVENTIONS

PCD offers a variety of musical dimensions as summarized in Table 4. These dimensions, referred to as ComposerID, WorkNo, MeasRange, PID, VersionID, StemType, and Reverb, are used in the filenames of the provided WAV audio files. The ComposerID specifies the composer (see Figure 2a). WorkNo indicates the Opus, BWV, or KV number of the work, and MovementNo denotes the number of the movement from which the excerpt was selected. The MeasRange dimension specifies range of the excerpt in measures. PID identifies the performer, as introduced in **Section 4.1**, and VersionID the version. StemType refers to the post-processing configurations presented in **Section 4.6**. Reverb refers to the presence of artificial reverb added in the post-production. The audio filename using the instances in the **Example** column in Table 4 is `Bach_BWV1056-01-mm001-008_YO-V2_OP_reverb.wav`. It represents Bach's (ComposerID) Piano Concerto in F Minor, BWV 1056 (WorkNo), 1st Movement (MovementNo), Measures 1–8 (MeasRange), played by `YO` (PID), second version (VersionID), which includes piano part plus orchestral accompaniment (StemType) with artificial reverb (Reverb).

4.3 SYNCHRONIZATION

Similar to the development of other multitrack datasets, the PCD curation encountered several challenges regarding the alignment of separate tracks. The missing interaction between the performer and other musicians constitutes a key challenge in a multitrack recording setting. As Li et al. (2019) suggest, audio-visual cues may help the musicians when playing along with a given audio track. In the recording process of the PCD, we used only audio cues, which served as a guide for the performers alongside the pre-recorded orchestral accompaniments.

The main objective of PCD is to provide piano recordings, which are synchronous to the *original* backing tracks by MMO. This design choice enables the dataset's reproducibility while restricting the freedom of interpretations since the pianists *must* steadily adapt their tempo to the orchestral track. To overcome the

challenges posed by the recording settings, we provided metronome-like click tracks in the form of sonified measure and beat annotations. We first manually annotated the measure positions in the backing track where the orchestra is active. Note that piano concertos often involve relatively long piano solo passages. In these sections, we employed linear interpolation to estimate the measure positions. This approach guaranteed that the sonified metronome-like click tracks retained consistent tempo in sections where the backing track is silent.

For the beats, we initially experimented with manual annotations. However, we found that using manual annotations based on the orchestral accompaniments was ineffective for the performers, as the tempo changes within measures were often inconsistent. As an alternative, we again utilized linear interpolation to estimate the beats within the manually annotated measure positions based on the time signature of the piece. This approach resulted in equidistant beats interpolated between the manual measure annotations, which were more helpful for the pianists than manual beat annotations.

Only for the recordings of Rachmaninov's Piano Concerto No. 2 in C Minor, Op. 18, we adopted a more involved iterative approach for the generation of beat annotations. For example, piano-only sections meant to be played with rubato (rather than a consistent tempo) were annotated by the performer such that the click tracks would match the tempo fluctuations in their interpretation. This facilitated a more natural-sounding recording of solo sections with larger variations in tempo.

Finally, we sonified measure and beat positions with different frequencies to aid the pianists during the recording process. Depending on the preference of the musician, we either activated or deactivated the click tracks during recording to allow for more agogical playing.

4.4 RECORDING PROCESS

In this section, we outline the technical details about the recording process. The performances in rooms `R1`, `R3`, and `R4` were recorded using a stereo spot microphone setup with *Schoeps MK4* cardioid microphones placed near the bend of the grand piano body (see Figure 3). Comparable high-end microphones like the MK4 are often used in similar professional recording setups. The exact position of the microphones was individually adjusted to the acoustics of each recording space and the characteristics of the instrument, roughly following an *ORTF (Office de Radiodiffusion Télévision Française)* setup. The microphone signals were recorded using a *RME Babyface Pro FS* audio interface and the *REAPER*[4] digital audio workstation. Recordings were initially stored in the *WAV* format with a sampling rate of *44.1* kHz and *24* bits per sample. The orchestral accompaniment and

DIMENSION	DESCRIPTION	EXAMPLE
ComposerID	Composer Identifier	`Bach`
WorkNo	Op. / BWV / KV	`BWV1056`
MovementNo	Movement Number	`01`
MeasRange	Measure Range	`mm001-008`
PID	Performer Identifier	`YO`
VersionID	Version Identifier	`V2`
StemType	(O)rchestra / (P)iano	`OP`
Reverb	Presence of Reverb	`reverb`

Table 4 PCD dimensions, encoded in filenames.

Figure 3 An impression from the recording process in R4 with stereo spot microphones (two *Schoeps MK4*). To play synchronously with the orchestra, the performer listens to the MMO orchestral accompaniment (superimposed by click tracks) via *Beyerdynamic DT 770 Pro* headphones.

sonified click tracks were presented to the musicians via headphones (*Beyerdynamic DT 770 Pro*) and played back from the same *REAPER* session to ensure a synchronous recording of the piano part. The pianists had the possibility to record the movements in shorter segments and repeat individual sections, as is common in a studio recording process. This typically results in multiple takes for the same section, which are later edited to form a coherent performance (see **Section 4.6**). The performances in room R2 (the recordings of Rachmaninov's Piano Concerto No. 2 in C Minor, Op. 18) were captured in a similar fashion, only differing in the utilized equipment. These performances were recorded using a stereo pair of *Sennheiser MKH 8020* omnidirectional microphones in *AB* configuration with a spacing of *35* cm. The microphones were placed approximately 1 m from the bend of the grand piano body at the height of 145 cm. A *Steinberg UR22mkII* audio interface and the *Cubase*[5] digital audio workstation were used.

4.5 MMO PRE-PROCESSING

The backing tracks provided by MMO vary in recording quality and format. To provide consistent orchestral accompaniments suitable for recording the piano parts, we modified the original tracks in several ways.

First, some of the MMO recordings were supplied in multiple sections (e.g., including just one page of sheet music). To have backing tracks of the entire movements, we joined the audio files which belong to the same movement. The resulting backing tracks are single audio files that serve as a continuous reference timeline for the dataset. Furthermore, we removed audible waveform artifacts at the splitting points. Finally, we removed the silence at the beginning and end of CD audio files. Note that this results in a shorter total duration of the reference timeline than the sum of the MMO tracks. All timings in the provided annotations and documentation refer to the reference timeline of the backing tracks created in

this process. We conducted all the modifications with Python scripts, which allows for reproducing our backing tracks from original MMO files.

Second, we removed some clicks in the backing track, provided in MMO in pauses of the orchestral accompaniment where the pianist plays solo. In the rendered excerpts, all clicks are always deactivated. This applies to the backing tracks of `Bach_BWV1056-01`, `Beethoven_Op037-01`, `Beethoven_Op058-02`, `Mendelssohn_Op025-01`, `Mozart_KV414-01`, `Rachmaninoff_Op018-01`, `Rachmaninoff_Op018-02`, `Rachmaninoff_Op018-03`, and `Schumann_Op054-01`.

Third, we finally employed some additional cosmetic pre-processing, including the removal of background noises (using the *iZotope RX8 Audio Editor*) and equalization for more consistent timbral qualities between pieces.

4.6 POST-PRODUCTION

The post-production of the recorded performances was conducted in three steps. First, we edited the recorded takes in *REAPER* to create a coherent rendition of the piano part. The takes were chosen to reduce playing mistakes while still maintaining a consistent musical arc in the performance, similar to post-production in a recording studio. Note that we maintained the timeline of the backing track in the post-production. Only the piano recording was edited to achieve good synchronicity with the MMO orchestral accompaniments.

Second, equalization was applied to the piano recordings to ensure consistent timbral qualities within our dataset without overcompensating the differences between instruments and recording spaces. Some minor noise removal similar to the MMO pre-processing was necessary to remove background noises. Third, to increase the coherence between the piano part and orchestral accompaniment, we applied artificial reverberation to both tracks simultaneously using the *FabFilter Pro-R2* algorithmic reverb software. All tracks are available with and without artificial reverberation to facilitate different use cases (see below). For the dataset, the post-processed excerpts were exported as *WAV* files with 44.1 kHz sampling rate and 16 bits per sample in six different configurations:

- `OP_reverb`: Piano part plus orchestral accompaniment with artificial reverb
- `OP`: Piano part plus orchestral accompaniment without artificial reverb
- `P_reverb`: Piano part only with artificial reverb
- `P`: Piano part only without artificial reverb
- `O_reverb`: Orchestral accompaniment only with artificial reverb
- `O`: Orchestral accompaniment only without artificial reverb

During the recordings of `Beethoven_Op015-01`, and `Mozart_KV467-01`, the orchestral accompaniment was erroneously played back with a rate of 0.995 (`Beethoven_Op015-01`) and 1.005 (`Mozart_KV467-01`), which results in a slightly slower or faster piano part relative to the backing track with the original playback speed. This mistake was corrected in the post-production with *Elastique Pro v3.3.3* by applying time-scale modification (with rates 1.005458 and 0.994616, respectively).

5. PCD INTERFACES

The main motivation of PCD is to provide a freely available and well-documented multitrack dataset to support MIR research on orchestral music, particularly piano concertos. To this end, the dataset is made publicly accessible through different interfaces in order to support scientific exchange and ensure the reproducibility of scientific results.

Interactive interfaces can lower barriers to access datasets and research results. This can be achieved through features such as playback functionalities (Gasser et al., 2015; Röwenstrunk et al., 2015; Jeong et al., 2017). To provide an interactive medium for the researchers, we use an open-source audio player (Werner et al., 2017) integrated in a web interface, which allows the listener to switch between multiple audio tracks while synchronously indicating the playback position of the audio tracks. As default visualization, the interface offers an overview of the six configurations of stems, as presented in **Section 4.6**. The main page is subdivided into a section called *Excerpts*, which includes links to recorded piano concerto sections with a dedicated sub-page for each excerpt. Figure 4 shows a screenshot of an
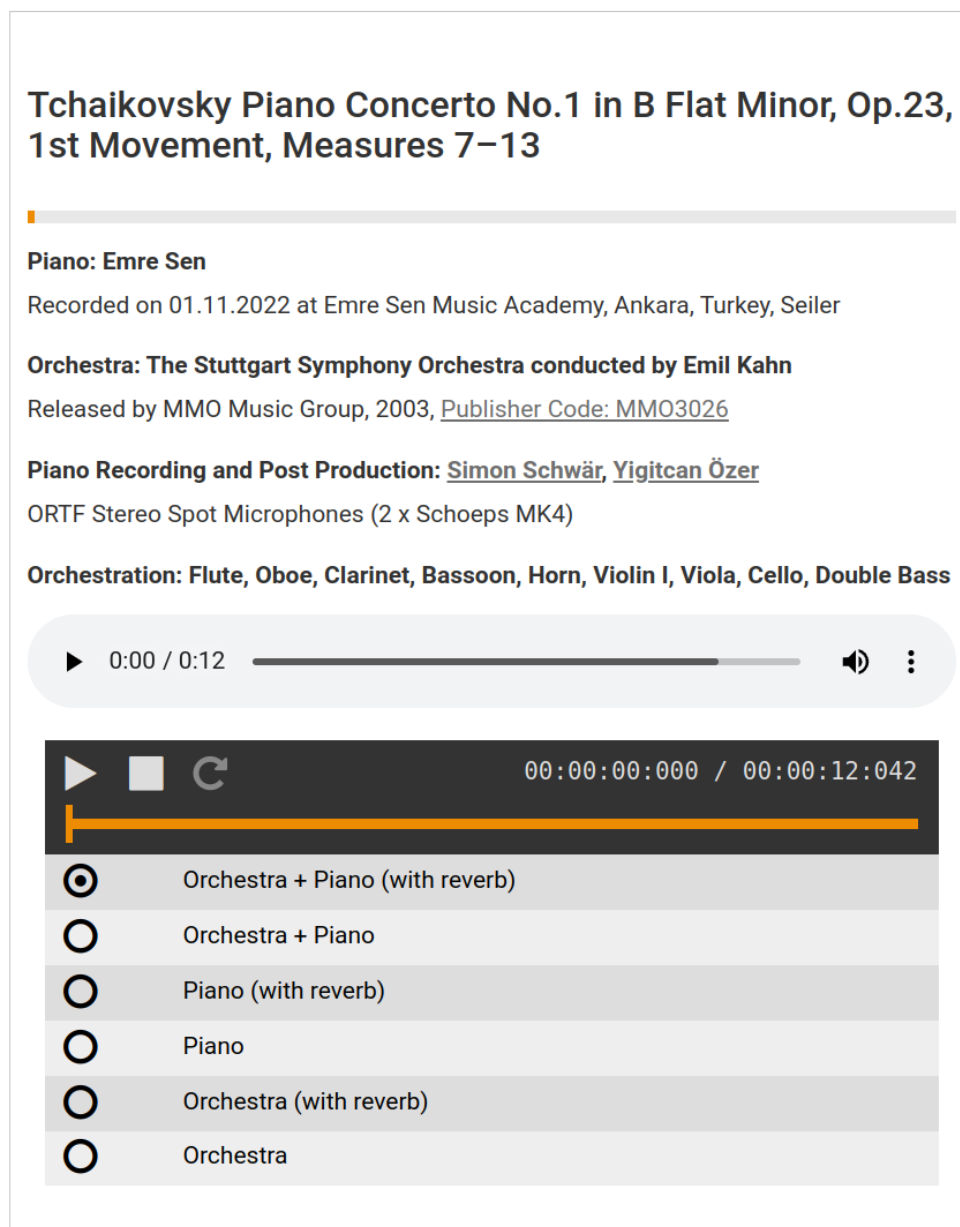

Figure 4 Screenshot of our web-based interface with Track Switcher (Werner et al., 2017), which comprises six tracks of dry and reverberant recordings of an excerpt from Tchaikovsky's Piano Concerto No. 1 in B Flat Minor, Op. 23, 1st movement.

exemplary sub-page,[6] which hosts the multitrack audio files for an excerpt selected from Tchaikovsky's Piano Concerto No. 1 in B Flat Minor, Op. 23, 1st movement.

6. APPLICATIONS TO MUSIC SOURCE SEPARATION

In this section, we highlight the potential of PCD by means of a case study in MSS. Here, we consider the separation of piano concertos into piano and orchestral tracks, which can be regarded as a lead-accompaniment separation task (Rafii et al., 2018). To this end, we use the pre-trained model by Özer and Müller (2022), which is a spectral-based U-Net architecture (Hennequin et al., 2020).[7] The training procedure of the pre-trained model is based on artificial random mixes of samples from the solo piano repertoire (e.g., piano sonatas, mazurkas) and orchestral pieces without piano (e.g., symphonies) to simulate piano concertos. While this method cannot simulate the harmonic and rhythmic relationships between different instruments in a real recording, it trains the model to identify the unique sound qualities of multiple musical sources.

For the quantitative and subjective evaluation of the pretrained MSS model by Özer and Müller (2022), we use the widely-used Signal to Distortion Ratio (SDR) (Vincent et al., 2006), computed with the *BSSEval*[8] Python library. Using dry and reverberant recordings, we assess the separation results along the various dimensions of PCD.

In Figure 5, we present the results based on reverberant recordings on the top and dry recordings on the bottom.

The trend generally indicates that SDR results based on dry recordings are better than those from reverberant recordings.

To get a first impression of the model's performance by different dimensions, Figure 5a provides an overview of the average SDR values per composer. Beethoven's piano concertos have the highest SDR value for separated piano. Note that Bach's piano concerto has the highest unison overlap between piano and orchestra, which results in a relatively lower separation performance for both parts.

For our next evaluation, we focus on the model's performance per performer (see Figure 5b). The SDR results reveal that piano separation outperforms orchestra separation for each performer. In particular, the excerpt played by LR yields the highest SDR values for both parts, both for dry and reverberant recordings. Note that the results for LR are based on a single excerpt, which is an easier passage for the model to separate. While separating piano and orchestra leads to similar SDR results for ES, MM, and YO, the piano separation is better than the orchestra by a wider margin for JL.

Figure 5c illustrates the comparison of SDR values across different acoustic environments. Similar to our previous analysis based on the SDR values per performer, piano separation yields higher SDR values compared to orchestra. The results indicate that the highest SDR value varies based on the artificial reverb in different acoustic conditions. Whereas the highest SDR value occurs in R2 for piano separation for reverberant recordings, the average SDR value for the dry recordings is the highest in R1.
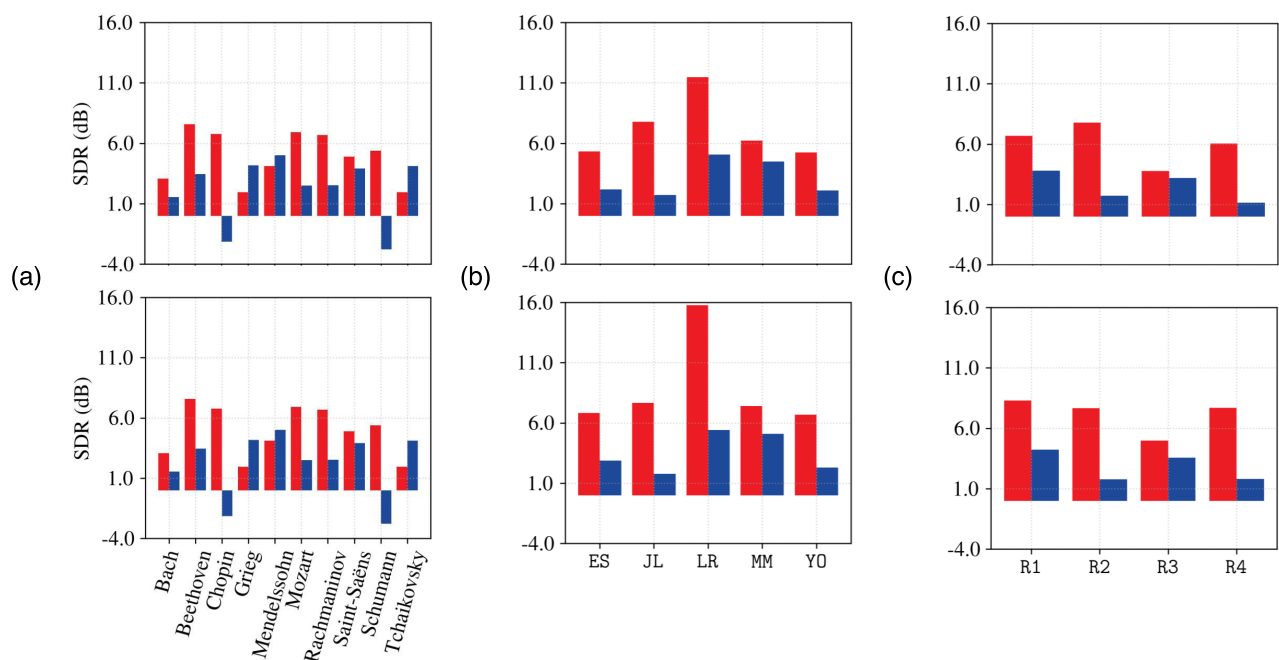

Figure 5 Comparison of Signal-to-Distortion-Ratio (SDR) values for the separation of piano (red) and orchestra (blue), averaged over **(a)** composer, **(b)** performer, and **(c)** acoustic environment. The bar plots on top indicate the results based on reverberant recordings, on the bottom dry.

For the future, our goal is to explore musically plausible data augmentation methods that simulate more realistic mixtures. To further enhance the separation performance, avenues of research may be to integrate a transcription model as proposed by Manilow et al. (2020) or to use the real and imaginary part of the STFT as an input to the network, following the Complex-as-Channel approach by Choi et al. (2020). Furthermore, we intend to investigate objective evaluation measures, e.g., 2f-score introduced by Kastner and Herre (2019), to assess the source separation performance of piano concertos.

7. CONCLUSION

In this paper, we introduced the Piano Concerto Dataset (PCD), which comprises excerpts from piano recordings and orchestral accompaniments of piano concertos ranging from the Baroque to the Post-Romantic era. Using backing tracks from the music publisher Music Minus One (MMO), we recorded 15 different piano concertos played by five performers with different instruments under varying acoustic conditions. To address the challenge of precise synchronization with pre-recorded orchestra accompaniments, we created click tracks to guide the pianists during the recording process. As a main contribution of PCD, we provide 81 excerpts of dry and reverberant recordings of piano and orchestra stems and their mixtures. We release the dataset via an interactive web-based interface to provide convenient access. Diverse musical dimensions of PCD enable various applications for MIR research, particularly for quantitative and subjective evaluation of source separation models.

NOTES

1 https://www.halleonard.com/series/MMONE.
2 https://www.mutopiaproject.org.
3 https://mirdata.readthedocs.io/en/latest/.
4 https://www.reaper.fm/.
5 https://www.steinberg.net/cubase/.
6 https://www.audiolabs-erlangen.de/resources/MIR/PCD/.
7 https://github.com/deezer/spleeter.
8 https://github.com/sigsep/bsseval.

ACKNOWLEDGEMENTS

FUNDING INFORMATION

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Yigitcan Özer is the main contributor to this article's recording process, dataset creation, and writing. Simon Schwär was the sound engineer responsible for the recording and post-production process. Emre Sen, Jeremy Lawrence, Yigitcan Özer, and Meinard Müller were the main performers of the dataset. Vlora Arifi-Müller contributed to the dataset curation. Meinard Müller closely supervised this work and contributed with Simon Schwär to the article's writing. All authors have read and agreed to the published version of the manuscript.

AUTHOR AFFILIATIONS

Yigitcan Özer orcid.org/0000-0003-2235-8655
International Audio Laboratories Erlangen, Germany
Simon Schwär orcid.org/0000-0001-5780-557X
International Audio Laboratories Erlangen, Germany
Vlora Arifi-Müller orcid.org/0009-0006-3477-7143
International Audio Laboratories Erlangen, Germany
Jeremy Lawrence orcid.org/0009-0007-6702-2190
International Audio Laboratories Erlangen, Germany
Emre Sen orcid.org/0000-0003-0115-8670
Emre Sen Music Academy, Ankara, Turkey
Meinard Müller orcid.org/0000-0001-6062-7524
International Audio Laboratories Erlangen, Germany

REFERENCES

Arzt, A., Widmer, G., and **Dixon, S.** (2008). Automatic page turning for musicians via real-time machine listening. In *Proceedings of the European Conference on Artificial Intelligence (ECAI)*, pages 241–245, Amsterdam, The Netherlands. IOS Press.

Bay, M., Ehmann, A. F., and **Downie, J. S.** (2009). Evaluation of multiple-f0 estimation and tracking systems. In *Proceedings of the International Society for Music*

Information Retrieval Conference (ISMIR), pages 315–320, Kobe, Japan.

Bittner, R. M., Fuentes, M., Rubinstein, D., Jansson, A., Choi, K., and **Kell, T.** (2019). mirdata: Software for reproducible usage of datasets. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 99–106, Delft, The Netherlands.

Bittner, R. M., Salamon, J., Tierney, M., Mauch, M., Cannam, C., and **Bello, J. P.** (2014). MedleyDB: A multitrack dataset for annotation-intensive MIR research. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 155–160, Taipei, Taiwan.

Bugler, A., Pardo, B., and **Seetharaman, P.** (2020). A study of transfer learning in music source separation. arXiv:2010.12650.

Cano, E., FitzGerald, D., Liutkus, A., Plumbley, M. D., and **Stöter, F.** (2019). Musical source separation: An introduction. *IEEE Signal Processing Magazine*, 36(1):31–40. DOI: https://doi.org/10.1109/MSP.2018.2874719

Caplin, W. E. (1998). *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*. Oxford University Press, USA.

Chen, K., Dong, H.-W., Luo, Y., McAuley, J., Berg-Kirkpatrick, T., Puckette, M., and **Dubnov, S.** (2022). Improving choral music separation through expressive synthesized data from sampled instruments. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, Bengaluru, India.

Chiu, C.-Y., Hsiao, W.-Y., Yeh, Y.-C., Yang, Y.-H., and **Su, A. W.-Y.** (2020). Mixing-specific data augmentation techniques for improved blind violin/piano source separation. In *Proceedings of the Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6. DOI: https://doi.org/10.1109/MMSP48831.2020.9287146

Choi, W., Kim, M., Chung, J., Lee, D., and **Jung, S.** (2020). Investigating U-Nets with various intermediate blocks for spectrogram-based singing voice separation. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, online.

Cole, W. (1997). *The Form of Music*. The Associated Board of the Royal Schools of Music (ABRSM), London, UK.

Cuesta, H., Gómez, E., Martorell, A., and **Loáiciga, F.** (2018). Analysis of intonation in unison choir singing. In *Proceedings of the International Conference of Music Perception and Cognition (ICMPC)*, pages 125–130, Graz, Austria.

Dannenberg, R. B. and **Raphael, C.** (2006). Music score alignment and computer accompaniment. *Communications of the ACM, Special Issue: Music Information Retrieval*, 49(8):38–43. DOI: https://doi.org/10.1145/1145287.1145311

Défossez, A. (2021). Hybrid spectrogram and waveform source separation. In *Proceedings of the ISMIR 2021 Workshop on Music Source Separation*, Online.

Ewert, S., Müller, M., and **Grosche, P.** (2009). High resolution audio synchronization using chroma onset features.

In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 1869–1872, Taipei, Taiwan. DOI: https://doi.org/10.1109/ICASSP.2009.4959972

Ewert, S., Pardo, B., Müller, M., and **Plumbley, M.** (2014). Score-informed source separation for musical audio recordings: An overview. *IEEE Signal Processing Magazine*, 31(3):116–124. DOI: https://doi.org/10.1109/MSP.2013.2296076

Gasser, M., Arzt, A., Gadermaier, T., Grachten, M., and **Widmer, G.** (2015). Classical music on the web –user interfaces and data representations. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 571–577, Málaga, Spain.

Girdlestone, C. M. (1948). *Mozart & His Piano Concertos*. Cassell & Company Ltd., London, UK.

Goto, M., Hashiguchi, H., Nishimura, T., and **Oka, R.** (2002). RWC music database: Popular, classical and jazz music databases. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 287–288, Paris, France.

Gover, M. and **Depalle, P.** (2020). Score-informed source separation of choral music. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*.

Hennequin, R., Khlif, A., Voituret, F., and **Moussallam, M.** (2020). Spleeter: A fast and efficient music source separation tool with pre-trained models. *Journal of Open Source Software*, 5(50):2154. DOI: https://doi.org/10.21105/joss.02154

Jansson, A., Humphrey, E. J., Montecchio, N., Bittner, R. M., Kumar, A., and **Weyde, T.** (2017). Singing voice separation with deep U-net convolutional networks. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 745–751, Suzhou, China.

Jeong, D., Kwon, T., Park, C., and **Nam, J.** (2017). PerformScore: Toward performance visualization with the score on the web browser. In *Demos and Late Breaking News of the International Society for Music Information Retrieval Conference (ISMIR)*, Suzhou, China.

Kastner, T. and **Herre, J.** (2019). An efficient model for estimating subjective quality of separated audio source signals. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 95–99, New Paltz, New York, USA. DOI: https://doi.org/10.1109/WASPAA.2019.8937179

Kim, M., Choi, W., Chung, J., Lee, D., and **Jung, S.** (2021). KUIELab-MDX-Net: A two-stream neural network for music demixing. In *Proceedings of the ISMIR 2021 Workshop on Music Source Separation*, Online.

Li, B., Liu, X., Dinesh, K., Duan, Z., and **Sharma, G.** (2019). Creating a multitrack classical music performance dataset for multimodal music analysis: Challenges, insights, and applications. *IEEE Transactions on Multimedia*, 21(2):522–535. DOI: https://doi.org/10.1109/TMM.2018.2856090

Manilow, E., Seetharaman, P., and **Pardo, B.** (2020). Simultaneous separation and transcription of mixtures

with multiple polyphonic and percussive instruments. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 771–775. DOI: https://doi.org/10.1109/ICASSP40776.2020.9054340

Martínez-Ramírez, A. M., Liao, W.-H., Fabbro, G., Uhlich, S., Nagashima, C., and **Mitsufuji, Y.** (2022). Automatic music mixing with deep learning and out-of-domain data. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, Bengaluru, India.

McLeod, A., Schramm, R., Steedman, M., and **Benetos, E.** (2017). Automatic transcription of polyphonic vocal music. *Applied Sciences*, 7(12). DOI: https://doi.org/10.3390/app7121285

Miron, M., Janer, J., and **Gómez, E.** (2017). Monaural score-informed source separation for classical music using convolutional neural networks. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 55–62, Suzhou, China.

Miron, M. and **Martorell, A.** (2017). Bach10 Sibelius Dataset. DOI: https://doi.org/10.5281/zenodo.321361

Özer, Y. and **Müller, M.** (2022). Source separation of piano concertos with test-time adaptation. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 493–500, Bengaluru, India.

Petermann, D., Chandna, P., Cuesta, H., Bonada, J., and **Gómez, E.** (2020). Deep learning based source separation applied to choir ensembles. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 733–739, Montreal, Canada.

Porter, A. (2012). Evaluating musical fingerprinting systems. Master's thesis, McGill University, Montreal, Canada.

Prätzlich, T., Driedger, J., and **Müller, M.** (2016). Memory-restricted multiscale dynamic time warping. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 569–573, Shanghai, China. DOI: https://doi.org/10.1109/ICASSP.2016.7471739

Rafii, Z., Liutkus, A., Stöter, F.-R., Mimilakis, S. I., and **Bittner, R.** (2017). The MUSDB18 corpus for music separation. DOI: https://doi.org/10.5281/zenodo.1117372

Rafii, Z., Liutkus, A., Stöter, F., Mimilakis, S. I., FitzGerald, D., and **Pardo, B.** (2018). An overview of lead and accompaniment separation in music. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(8):1307–1335. DOI: https://doi.org/10.1109/TASLP.2018.2825440

Rosenzweig, S., Cuesta, H., Weiß, C., Scherbaum, F., Gómez, E., and **Müller, M.** (2020). Dagstuhl ChoirSet: A multitrack dataset for MIR research on choral singing. *Transactions of the International Society for Music Information Retrieval (TISMIR)*, 3(1):98–110. DOI: https://doi.org/10.5334/tismir.48

Röwenstrunk, D., Prätzlich, T., Betzwieser, T., Müller, M., Szwillus, G., and **Veit, J.** (2015). Das Gesamtkunstwerk Oper aus Datensicht – Aspekte des Umgangs mit einer heterogenen Datenlage im BMBF-Projekt "Freischutz Digital". *Datenbank-Spektrum*, 15(1):65–72. DOI: https://doi.org/10.1007/s13222-015-0179-0

Sarkar, S., Benetos, E., and **Sandler, M.** (2022). EnsembleSet: A new high quality dataset for chamber ensemble separation. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, Bengaluru, India.

Schedl, M., Hauger, D., Tkalčič, M., Melenhorst, M., and **Liem, C. C. S.** (2016). A dataset of multimedia material about classical music: PHENICX-SMM. In *Proceedings of the International Workshop on Content-Based Multimedia Indexing (CBMI)*, pages 1–4. DOI: https://doi.org/10.1109/CBMI.2016.7500240

Schering, A. (1905). *Geschichte des Instrumentalkonzerts*. Breitkopf & Härtel, Leipzig, Germany.

Schramm, R. and **Benetos, E.** (2017). Automatic transcription of a cappella recordings from multiple singers. In *Proceedings of the AES International Conference on Semantic Audio*, pages 108–115, Erlangen, Germany.

Series, B. (2014). Method for the subjective assessment of intermediate quality level of audio systems. *International Telecommunication Union Radiocommunication Assembly*.

Stoller, D., Ewert, S., and **Dixon, S.** (2018). Wave-Unet: A multi-scale neural network for end-to-end audio source separation. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 334–340, Paris, France.

Stöter, F., Uhlich, S., Liutkus, A., and **Mitsufuji, Y.** (2019). Open-Unmix – A reference implementation for music source separation. *Journal of Open Source Software*, 4(41). DOI: https://doi.org/10.21105/joss.01667

Takahashi, N. and **Mitsufuji, Y.** (2017). Multi-scale multi-band densenets for audio source separation. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 21–25. DOI: https://doi.org/10.1109/WASPAA.2017.8169987

Vincent, E., Gribonval, R., and **Févotte, C.** (2006). Performance measurement in blind audio source separation. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(4):1462–1469. DOI: https://doi.org/10.1109/TSA.2005.858005

Werner, N., Balke, S., Stöter, F.-R., Müller, M., and **Edler, B.** (2017). trackswitch.js: A versatile webbased audio player for presenting scientific results. In *Proceedings of the Web Audio Conference (WAC)*, London, UK.

Transactions of the International Society for Music Information Retrieval is a peer-reviewed open access journal published by Ubiquity Press.