Acoustic-digital hybrid synthesizer

Levin Schnabel
lschna23@student.aau.dk
Aalborg University
Copenhagen, Denmark

Dan Overholt
dano@create.aau.dk
Aalborg University
Copenhagen, Denmark

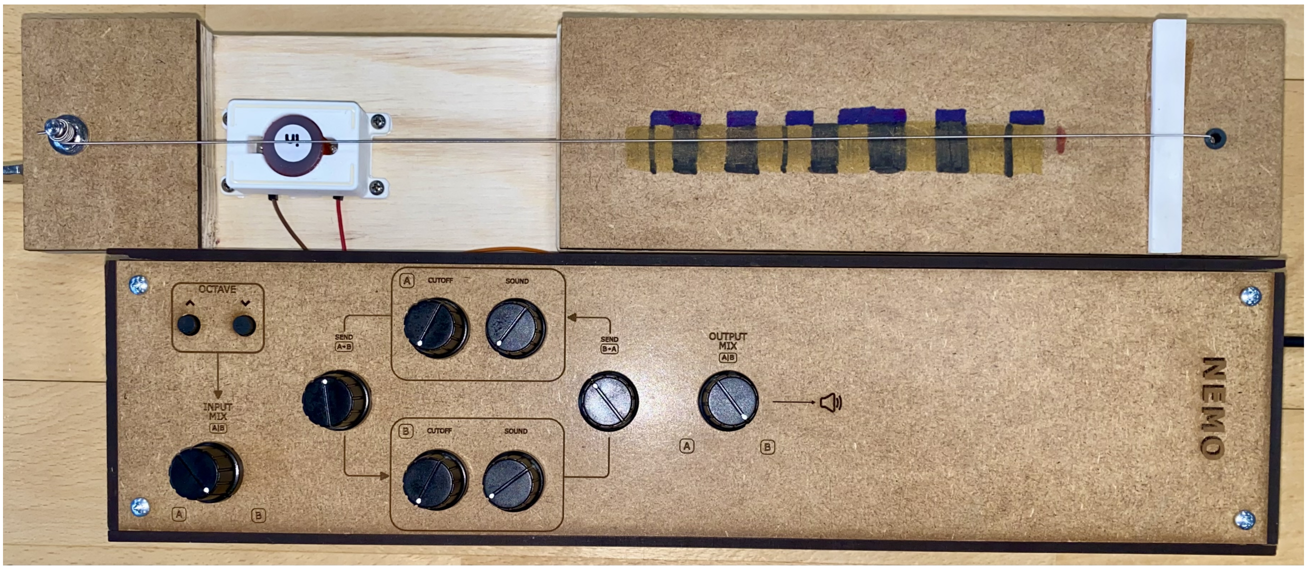

Figure 1: Prototype of an acoustic-digital synthesizer, 2024

Abstract

This paper explores the design and evaluation of an acoustic-digital hybrid instrument that aims to address key criticisms of Digital Musical Instruments (DMIs), particularly the separation of control and sound generation. By integrating an interactable physical string with coupled Finite Difference Schemes (FDS) for physical modeling synthesis, the instrument creates a tactile and responsive playing experience.

The instrument was evaluated through a mixed-methods approach, combining qualitative think-aloud protocols with the Musician's Perception of the Experiential Quality of Musical Instruments Questionnaire (MPX-Q). Results indicate that the instrument fosters curiosity and creativity but highlights challenges in achieving traditional acoustic playability, such as latency and perceptual dissonance. These findings emphasize the potential and limitations of acoustic-digital hybrids in reuniting control and sound, offering valuable insights for future developments in musical interface design.

Keywords

Acoustic-digital hybrid instruments, Physical modeling synthesis, Controller-generator paradigm

NIME '25, June 24–27, 2025, Canberra, Australia

1 Introduction

Within New Interfaces for Musical Expression (NIME) research a common research goal is to create expressive instruments [37]. However, many resulting Digital Musical Instruments (DMIs) are based on the controller/generator paradigm, which separates the instrument into two distinct components: one for controlling sound and the other for generating it [10]. That approach is in stark contrast to traditional acoustic instruments where "control" and sound are inherently linked [22]. While this approach has it's merits, it bears the risk of transforming the intricate process of designing an instrument into designing a user experience and turning musicians into system operators [18].

This shift can diminish the embodied, tactile, and emergent qualities of musical performance, where sound production and physical interaction are deeply intertwined. Musicians may find themselves managing parameters rather than engaging in a dynamic relationship with the instrument. In response to this challenge, the presented work investigates how acoustic-digital hybrid instruments can re-entangle control and sound generation, fostering richer and more intuitive musical interactions.

1.1 Shortcomings of the Controller/Generator Paradigm

Perry Cook highlights the need to "remutualize" the controller and generator in DMIs, pointing to a lack of intimacy between musician and instrument [10]. This critique has been further elaborated in subsequent research, identifying several key shortcomings of current DMIs:

- **Lack of passive haptic feedback:** Unlike traditional acoustic instruments, many DMIs fail to provide tactile

feedback that naturally arises from the physical interaction with an instrument [8, 10, 23].
- **Absence of sonority and 'ergotic' sounds [7]:** In acoustic instruments, controlling actions, and sound production are energetically linked, creating a direct and embodied relationship between the performer and the resulting sound. This energetic coupling is often missing in DMIs [10, 17, 22].
- **Discrete rather than continuous gesture mapping:** Musical gestures in DMIs are frequently treated as discrete commands or control-rate signals rather than as continuous, time-dependent functions, reducing the expressive potential of performance [10, 37].

These critiques emphasize how the separation of controller and generator can undermine intimacy in musical expression with a DMI.

1.2 Acoustic-digital hybrid instruments

In response to some of the abovementioned critiques, Michon et al. introduced the notion of acoustically driven hybrid instruments [24]. These use audio-rate signals (sampled at 44.1 kHz and above) capturing vibrations of acoustic objects such as strings or plates to excite "virtual" elements. Here, the notion of virtual usually refers to the usage of physical modeling synthesis. For brevity, the term acoustic-digital hybrid will be used to refer to this foundational concept.

1.2.1 Example instruments. There are plenty of examples showcasing acoustic-digital hybrid instruments. To name a few:

- Building on their conceptual foundation, Michon et al. introduced 3D-printed acoustic extensions that integrate with smartphones to enable more cohesive interactions with DMIs. Combined with physical models, these designs integrated the sound generation with various physical elements measurable with smartphones. An example is an ocarina driven by blowing into the phone's microphone [24].
- Davison et al. presented a "Self-Sensing Haptic Actuator for Tactile Interaction with Physical Modelling Synthesis" [11]. The haptic device simultaneously senses user input and provides haptic feedback using a single moving coil transducer. This bi-directional approach drives and is driven by physical modeling synthesis. They showcased the example of two connected plates.
- The Kalichord [29] drives virtual resonators modeled using waveguides with an audio-rate signal generated from a piezo microphone. While the authors classify it as an electro-acoustic instrument, the notion of acoustically driven hybrid instruments also fits.
- Lastly, the Strummi [19] is a string-based tabletop instrument. Similarly to the Kalichord, piezo microphones pick up the string's vibration, which serves as the excitation signal for a Karplus-Strong model.

These designs place the interaction with a physical model at the center. Hybridization is approached through high-fidelity synthesis, excited in real time by real-world signals. While this is the approach followed in this project, it is not the only way to address the shortcomings discussed in 1.1. Other notable methods include:

- **Feedback-based hybrid instruments**: Some instruments create a hybrid by closing a feedback loop between acoustic resonators and digital processing, using various sensing and actuation techniques. The physical body serves as both resonator and interaction surface, while digital systems manage and sometimes deliberately destabilize the loop. This creates complex, emergent behaviors. A prominent example is the *halldorophone* [33], a cello-like instrument used in numerous compositions.
- **Actively controlled instruments**: These instruments digitally shape the vibrational behavior of acoustic objects. Feedback control can be used to dampen, sustain, or otherwise modify specific vibrational modes, effectively altering the instrument's timbre in real time [4]. Such systems have also been applied as physical audio effects [3] or for highly detailed control of wave propagation within an instrument [14].

1.2.2 From the viewpoint of entanglement. The presented project was developed in the context of NIME2025's conference theme, "entangled NIME"[1]. Viewed through this lens, acoustic-digital hybrid instruments offer new possibilities for deepening the entanglement between human performers and digital sound synthesis through acoustic media. More broadly, they blur traditional boundaries between what is considered acoustic or digital performance, suggesting that these categories are no longer discrete but increasingly intertwined. While purely acoustic or digital instruments each have distinct merits, examples such as the *halldorophone* demonstrate the rich creative potential that emerges when these domains are meaningfully combined. Similarly, this project explores how real-world physical interactions can be tightly integrated with digital synthesis, encouraging a more complex, reciprocal relationship between musician, digital instrument, and sound.

1.3 Addressing Controller/Generator Shortcomings

Conceptually, acoustic-digital hybrid instruments have the potential to address the shortcomings outlined in Section 1.1. By incorporating an acoustic element that users physically interact with, these instruments naturally provide passive haptic feedback. The missing energetic link between action and sound can be mitigated by sensing the physical object at audio rates and using the signal as an excitation signal of a physical model. Additionally, this inherently enables a continuous interaction, overcoming the issue of discrete gesture interpretation. This project presents a case study of an acoustic-digital hybrid instrument and evaluates how effectively such instruments can overcome the controller/generator division.

2 Prototype implementation

This section details the design and creation of a prototypical instrument implementing the conclusions drawn in the previous section. Furthermore, insights gained through the continuous reflection of the design decision are shared. While informal and anecdotal, these design insights may guide future prototypical designs. For a formal prototype evaluation refer to section 3.
The implementation process resulted in the creation of two artifacts:

[1]NIME2025 conference webpage: https://nime2025.org/

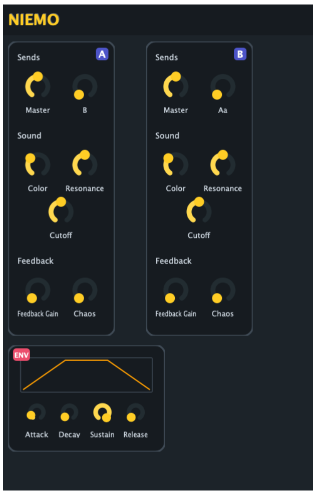

Figure 2: Screenshot of Artifact 1: A GUI-based software prototype.

(1) A software-only prototype (see figure 2) exploring the sound space of finite difference schemes (FDSs) with varying physical parameters, model coupling configurations, and parameters for the dynamic feedback algorithm. While constricting to regular computer-based hardware and graphical user interfaces (GUIs) favors the controller/generator paradigm, it also allows quick iterations. This enabled a deep exploration of the underlying sound space.
(2) The physical prototype of the final instrument is mainly concerned with the acoustic elements of this project. These include experiments mainly on sensing and inducing vibrations in a string, as well as a robust construction that can resist the stress caused by the string under tension.

The first artifact has been implemented as a VST plugin using the JUCE framework[2], whereas the final prototype uses the Bela platform[3]. The most severe difference between both environments is Bela's reduced amount of computing resources. As both platforms can be used with the C++ programming language, migrating code artifacts mainly involved replacing JUCE framework features with standard C++ library functions.

2.1 Software components

The core components of the system are dynamic FDSs and a mechanism to couple them flexibly [38, 39], the effort to compress (ETC) algorithm as an implementation of complexity-controlled feedback gain regularization [21], the Yin algorithm for fundamental frequency (F0) detection [12], and a phase vocoder for pitch-shifting signals [28].

2.1.1 Physical Modeling. All instruments showcased in Section 1.2.1 use physical modeling synthesis for their virtual components. Unlike synthesis techniques such as frequency modulation (FM) synthesis, which focuses on modeling the perceived sound, physical modeling simulates the behavior of vibrating objects, often referred to as resonators. These models are derived from mathematical formulations of real-world physical phenomena, and their parameters—such as string stiffness or material properties—are grounded in physics [35]. This makes physical modeling

[2]JUCE homepage: https://juce.com/
[3]Bela homepage: https://bela.io/

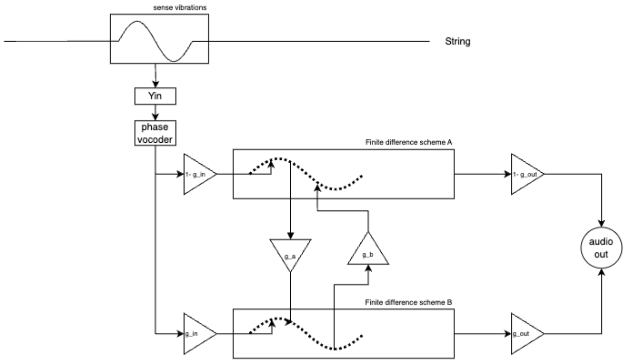

Figure 3: Block diagram of the final system

a natural and logical choice for integration with the acoustic input of the prototypical instrument.
Several approaches exist for implementing physical modeling. Finite-Difference Time-Domain (FDTD) methods, for example, are computationally intensive but offer greater flexibility in simulating complex systems [38]. In the development of the first artifact, FDTD methods could generate rich and promising soundscapes, as they enable precise control over how strongly different resonators interact. This flexibility and the accuracy of the models ultimately justified their use over other, more computationally efficient techniques.
Finite Difference Schemes (FDS) discretize both the spatial and temporal domains of a continuous system described by a partial differential equation (PDE). The spatial domain is represented as a grid u of finite points, where the system is modeled in discrete intervals. Each grid point u_n^l reflects the displacement at some position l and time n, with derivatives approximated through differences between neighboring grid values [6]. To use FDSs in real-time, the FDS needs to be resolved for u_{n+1}^l so that the next state of each point can be calculated iteratively in dependence on its own and its neighbor's past and present state [39].
The grid size of a FDS is determined by the physical parameters to ensure numerical stability. However, dynamically changing these parameters requires the grid to adapt accordingly. To maintain stability and minimize artifacts during such changes, Willemsen et al. proposed a method for smoothly adding or removing grid points [38]. This is achieved by introducing a second virtual grid at the boundaries of the "real" grid. The virtual grid enables smooth interpolation of displacements between the two grids while new grid points are incrementally added or removed.
The final system consists of two coupled FDSs, where the coupling force in either direction can be configured by the user. In the first artifact, each scheme was also coupled to itself, introducing additional feedback paths that further enriched the interaction dynamics. The models can be excited by any audio signal by updating the displacement at a specified grid point to reflect the sum of the current displacement and the incoming audio signal. See figure 3 for an overview of the system setup.

2.1.2 Dynamic feedback gain control. The introduction of a feedback path within an instrument opens up a vast new sonic space, offering numerous possibilities for creative sound design. Feedback can be highly rewarding and expressive when it enhances the dynamics of a signal but becomes frustrating when it locks into saturating feedback, reducing the agency of the musician. To address this, Kiefer et al. proposed an algorithm to automatically adjust feedback-contributing parameters based on the signal's complexity [21].

At the core of their approach is a complexity metric called Effort to Compress (ETC), designed to measure the dynamics of the signal. A low ETC value indicates the presence of saturating feedback, characterized by a dominant resonant frequency and reduced signal variability. ETC is computed as the number of iterations required to losslessly compress a sequence of symbols, where symbols represent quantized features of the audio signal, such as root mean square loudness (RMS). This process is based on the Non-Sequential Recursive Pair Substitution algorithm, which recursively replaces the most frequent pair of symbols with a new symbol until the sequence becomes constant.

Using ETC, the feedback gain can be dynamically controlled such that the gain is reduced when complexity drops, counteracting the buildup of saturating positive feedback. Optional smoothing and damping parameters in the gain control function allow users to fine-tune the responsiveness and effects of the gain management system, providing an intuitive way to maintain the instrument's expressiveness without directly manipulating the feedback gain. Furthermore, this frees the musician from constantly monitoring the feedback and leaves room to interact with the feedback system in deeper ways.

The proposed algorithm was implemented in the software-only prototype where each FDS was coupled to itself in a unidirectional manner. The force of this coupling was controlled using the gain management system outlined earlier. Users could adjust key parameters, including the maximum feedback gain and a "chaos" parameter, which reduced the dampening of the gain control as its value increased. These parameters introduced intriguing dynamic behaviors, enabling the creation of unique timbres that were otherwise not achieved. Notably, the interaction between two coupled FDSs, each with its own feedback path, allowed for a wide range of combinations, resulting in a highly complex and rich system.

While the algorithm was tested on the Bela platform, its recursive nature imposed significant computational demands. Without prior optimization, the algorithm could not be run in conjunction with other processing tasks. Chris Kiefer later evaluated the computational requirements of various complexity metrics and their effectiveness in detecting saturating feedback [20]. Among the alternatives, the random projection complexity metric presented a less resource-intensive option compared to ETC, as it relies on lossy compression. However, it requires more parameter tuning and was ultimately not adopted for the second artifact due to time constraints.

2.1.3 Yin algorithm for F0-detection. By tracking the fundamental frequency of the physical string signal, the parameters of the FDS can be adjusted to resonate at this frequency. This allows the perceived pitch of the FDS to be aligned with the pitch of the input string signal. The YIN algorithm, developed by De Cheveigné and Kawahara [12], is an autocorrelation-based single F0 estimator. It has been chosen for its computational and conceptual simplicity. This section describes the key steps of the YIN algorithm as presented in their work.

Instead of relying on the autocorrelation function, YIN uses the difference function (DF) to minimize errors, particularly those caused by amplitude variations. The DF of a signal shifted by the lag τ is defined as

$$d_t(\tau) = \sum_{j=1}^{W} (x_j - x_{j+\tau})^2 \tag{1}$$

where x_t is assumed to be a periodic signal with the period T and W is the length of the windowed signal.

The difference function is then normalized by its cumulative mean to avoid errors at low lags and to provide stable results even for pseudo-periodic signals, as described in the original paper:

$$d'_t(\tau) = \begin{cases} 1, & \text{if } \tau = 0 \\ \frac{d_t(\tau)}{\frac{1}{\tau}\sum_{j=1}^{\tau} d_t(j)}, & \text{otherwise} \end{cases} \tag{2}$$

To mitigate subharmonic errors, a threshold value of 0.1 is applied. The first candidate τ for which $d'_t(\tau)$ falls below this threshold is further refined by increasing τ until it is a local minimum of d'_t. The confidence of the period estimate is then quantified by the following measure:

$$\text{probability} = 1 - d'_t(\tau) \tag{3}$$

as smaller normalized differences between the original and shifted signal indicate higher periodicity. Finally, the initial period estimate τ is refined using parabolic interpolation improving the accuracy of the pitch detection.

This F0-tracking method can be further enhanced by differentiating between slowly and rapidly varying F0 components. A Kalman filter, as proposed by Christensen [9], could enable the system to distinguish between sustained notes and dynamic gestures such as vibratos and slides.

Latency is a critical factor in acoustic-digital instruments. Since F0 tracking relies on the analysis of at least one full period, it introduces an inherent latency that may impact the musical intimacy of performance [31]. Although this latency could potentially disrupt the sense of immediacy in musical interaction, priority was given to the player's ability to maintain tonal control over the instrument's output.

The final prototype uses a window with $W = 512$ samples, creating an inherent delay of 11.61 milliseconds. The algorithm can theoretically track periods of $f_{min} = \frac{44100}{512} \approx 86.13Hz$ assuming a standard sampling frequency. For reference, the lowest playable note on a guitar with standard tuning is about 83 Hz. For higher accuracy, it is advised to choose a window size able to accommodate two or more periods of the lowest F0 to detect.

2.1.4 Pitch shifting. In a traditional guitar, fret spacing follows the principles of equal temperament tuning, where each semitone corresponds to a fixed frequency ratio ($r = 2^{1/12}$). This results in an exponential relationship between string length and pitch, leading to progressively smaller fret spacing as the pitch increases. By shifting the pitch of the physical input signal, the dependency between the string's perceived pitch and the FDS is removed. With the FDS's excitation signal arbitrarily adjustable, this constraint is eliminated, allowing for equally spaced frets while maintaining accurate pitch intervals.

To enable a fretboard with equally spaced frets, the tracked F0 (as described in 2.1.3) is mapped to the corresponding semitone's F0. This requires an initial calibration step, where the fundamental frequency $f_{0_{L_0}}$ is measured at a known string length L_0. The string length L is assumed to be inversely proportional to the measured frequency f_0, expressed as:

$$L \propto \frac{1}{f_0}, \quad \text{or equivalently:} \quad L = L_0 \cdot \frac{f_{0_{L_0}}}{f_0}. \tag{4}$$

This formulation neglects the effects of tension but remains sufficiently accurate for practical use. To address inaccuracies caused by lower tension in the open string calibration ($f_{0_{L_0}}$), a fretted position, where tension is closer to typical playing conditions, is selected as the calibration reference. Knowing L, a semitone number can be calculated by:

$$semitone = semitone_{max} - ((L_0 - L)/L_{step}), \tag{5}$$

where $semitone_{max}$ is highest possible semitone number and L_{step} is the desired distance between semitones. Rounding this value leads to quantized semitones. The final semitone is blended with the rounded semitone to increase playability while keeping keeping the pitch a continuous function of length for expressivity. The frequency to transpose to can then be calculated by:

$$f_{new} = f_{ref} * 2^{semitone/12} \tag{6}$$

where f_{ref} defines which frequency corresponds to the first semitone.

The pitch-shifting algorithm is implemented using the phase vocoder, as described by Mark Dolson [13]. Phase vocoder-based effects typically involve an analysis stage, where the signal is transformed into the frequency domain using the Short-Time Fourier Transform (STFT) applied to overlapping windows. Signal manipulations are performed in the frequency domain by modifying the amplitude and phase of the spectral components. In the synthesis stage, the time-domain signal is reconstructed via the inverse Fourier transform and the overlap-add (OLA) method, ensuring continuity between overlapping windows [28].

In the analysis stage, each overlapping window is transformed into the frequency domain using the Fourier transform. The phase information of each spectral bin is tracked over successive analysis frames, allowing the instantaneous frequency to be computed as:

$$\Delta\phi(k, n) = \phi(k, n) - \phi(k, n-1) - 2\pi k \frac{h_a}{N} \tag{7}$$

where $\phi(k, n)$ is the phase of bin k in frame n, N is the FFT size, and h_a is the analysis hop-size.

To pitch shift the input signal by a factor R, the synthesis hop size h_s is set to equal $R * h_a$. The phase increments are adjusted proportionally so that

$$\phi_{\text{syn}}(k, m) = \phi_{\text{syn}}(k, m-1) + \Delta\phi(k, n) \cdot \frac{h_s}{h_a} \tag{8}$$

After phase and amplitude adjustments, the signal is reconstructed in the time domain using the inverse Fourier transform. Overlapping frames are combined using the overlap-add method. Finally, linear interpolation is applied to ensure that the output matches the original signal duration, effectively compressing or stretching the time-scaled output to achieve the pitch shift.

2.2 Hardware

The instrument's hardware primarily consists of an interactable string, an electromagnetic pickup, and traditional control elements such as potentiometers. The enclosure was laser-cut and reinforced with a plywood base positioned beneath the string to enhance structural integrity. The electromagnetic pickup coil is housed within a 3D-printed enclosure for mounting.

2.2.1 Sensing and actuating the string. Similar to an electric guitar, the string's vibrations can be detected using conventional guitar pickups or a coil with a $1.9k\Omega$ impedance and a metal core with an attached magnet. They are compatible with the Bela platform's audio input without any additional amplification.

Conversely, string actuation can be achieved by applying an electromagnetic field modulated by an audio signal. This technique has been studied and refined in prior research [2, 5], embraced by the DIY community[4], and implemented in commercial products such as the E-bow[5] [16].

To explore string actuation on the physical instrument, multiple experiments were conducted using a setup similar to that employed for signal pickup. However, the actuation setup utilized a 15Ω impedance coil from a 12V direct current relay. The input signal was amplified via a 12V audio amplifier circuit and sent to the output coil, forming a feedback loop through the string. While the system successfully induced vibrations in the open string, it failed to do so when the string was fretted. This behavior may align with the observations discussed in Section 2.1.4, where the imperfect construction caused a significant increase in string tension when fretted.

An informal yet critical learning from these experiments is that acoustic-digital hybrid instruments demand precise construction of all acoustic elements. In this case, the short string length limited sustain, impeding the output coil to maintain vibration effectively.

3 Evaluation

Evaluating an instrument designed to blur the boundaries between controller and generator requires a holistic approach. Traditional human-computer interaction (HCI) evaluation methods, such as task-based approaches that emphasize controllability [36], appear inadequate in this context [1]. However, evaluating the instrument as a whole provides only a case study with limited generalizability. Additionally, assessing musical intimacy poses a challenge, as experiment participants often have insufficient time to develop a meaningful relationship with a novel instrument. This approach risks disregarding the evolving interaction between musician and instrument [27].

To address these challenges, this evaluation focuses on specific aspects of the instrument, aiming to produce more generalizable insights into the creation of remutualized instruments through acoustic-digital hybrids.

3.1 Methodology

The experiment employed a mixed-methods approach to evaluate the instrument. First, participants were asked to think aloud during their initial interaction with the instrument prototype. Beyond the initial calibration described in Section 2.1.4, no prior demonstration or instructions were provided on how to use the instrument. This qualitative approach aimed to capture participants' first impressions and intuitions, assessing the ability of acoustic-digital instruments to be intuitively understood and used in diverse ways to produce sound. While think-aloud protocols are sometimes criticized for interfering with musical tasks due to the cognitive demands of verbalization, they provide unfiltered insights into participants' initial thoughts and reactions [32].

[4]Open hardware DIY sustainer (accessed 14.12.2024): https://bitbucket.org/metalmarshmallow/mm-diy-sustainer/src/main/

[5]The E-bow, a commercial sustainer (accessed 14.12.2024): https://www.ebow.com/

Participants were given up to 10 minutes to explore the instrument, with the option to stop earlier if desired. This limited timeframe was intentionally chosen, reflecting the instrument's early development stage and the expectation that participants would not be able to acquire significant skill within a single session. Rather than aiming to assess expressive mastery, the focus was placed on first impressions and discoverability. The actual time spent engaging with the instrument was recorded as an additional indicator of intuitive accessibility. During exploration, the researcher observed participants' behavior and recorded their verbalizations (with prior consent) to capture all nuances for later thematic analysis.
Secondly, Gian-Marco Schmid's Musician's Perception of the Experiential Quality of Musical Instruments Questionnaire (MPX-Q) [30] was used as a quantitative evaluation method. This psychometrically validated questionnaire explores three dimensions:

- Experienced Freedom and Possibilities,
- Perceived Control and Comfort, and
- Perceived Stability, Sound Quality, and Aesthetics.

The MPX-Q facilitates comparison with other instruments pursuing similar design goals while offering a structured and comprehensive exploration of hybrid acoustic-digital instruments from a musician's perspective.
Before completing the questionnaire on a prepared computer via the Internet, participants were asked to engage musically with the instrument. The definition of "musical" was intentionally left open to encourage participants to explore the instrument without being nudged toward specific features. However, for less musically trained participants, blue markings on the fretboard corresponding to the C Phrygian scale were provided as a guide. Participants were instructed to include an introduction and an outro in their improvisation, offering minimal constraints while ensuring a basic structure to their musical exploration. The statements from the MPX-Q were rated on a 7-point Likert scale.

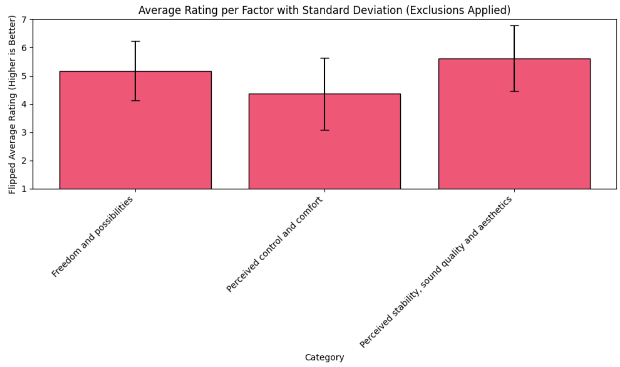

Figure 4: Average score and standard deviation per factor defined in MPX-Q. Each factor contains multiple categories.

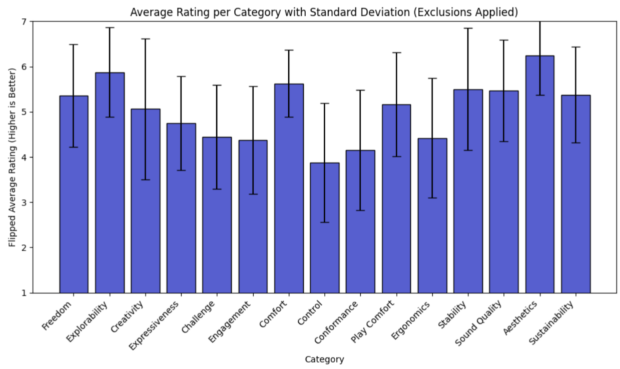

Figure 5: Average score and standard deviation per category defined in MPX-Q. Each category contains one or more questions

3.2 Participants

Participants were recruited from the Sound and Music Master's program at Aalborg University. Although the instrument is not explicitly designed for students from audio-related fields, this participant pool offered advantages: prior experience with novel musical interactions, familiarity with NIME concepts, and the ability to articulate experiences critically. However, this relatively homogeneous group could also introduce shared biases and assumptions about DMIs, likley limiting the generalizability of the findings. Future studies would benefit from targeting more diverse groups, including non-specialist musicians or performers outside academic settings, to better align evaluation contexts with real-world usage.

4 Results

In total, eight students participated in the experiment, aged between 23 and 34 years. All but one participant played an instrument recreationally, with 50% having semi-professional experience. Only one participant played a stringed instrument. The average initial time spent with the instrument was 8.5 minutes.

4.1 Discoverability and Intuition

Statements from the think-aloud protocol include quotes related to technical difficulties, confusion, and the ease or difficulty of understanding the instrument's functionality. Seven participants expressed confusion about the effect of certain dials and a lack of control, with statements like, "I have no clue what things are doing." This confusion extended to retrospective reflections on previous sound outcomes.
An aspect of exploration and discoverability was present as well. Although participants struggled to explain musical outcomes, they frequently expressed positive surprise: "Wow, I think I found something", "this is interesting", or "I don't know what's happening, but I like it".
This theme is reflected in the survey outcomes as well. The participants' difficulty in understanding the instrument's controls aligns with a lower average rating for the factor "Perceived control and comfort" (4.4), indicating challenges in achieving precision (see figure 4). The theme of a positive surprise in exploration is supported by the relatively high rating for Explorability (5.8), suggesting the instrument fosters curiosity despite initial confusion (see figure 5).

4.2 Interaction with the string

Many statements related to the string, its interaction with the instrument, and responsiveness. Participants approached the string differently from the traditional controls, often experimenting with plucking and fretting techniques rather than focusing on parameter adjustments.
Tactility was recurring in statements like, "I don't know if the

pitch difference is because of where I'm pressing on the string or if it's the way that I'm plucking it," and "I'm trying different picking positions." Participants attributed unexpected sounds more to their interaction with the string than to parameter setups.
The ratings for the categories expressiveness (5.0), challenge (4.6), and engagement (4.5) are surprisingly low when compared to the other ratings. Given the participant's statements regarding trying out various interaction techniques with the string, higher values of expressiveness and engagement were expected. These lower ratings could stem from participants' unfamiliarity with stringed instruments, limiting their ability to engage expressively with the string. Furthermore, a higher value of challenge was expected. Possibly, the participants felt too little impact on the sound when playing the string with various techniques. Hence, lacking challenges when trying to master a specific sound outcome.

4.3 Aesthetic reactions

Aesthetic reactions include emotional or descriptive responses to the sounds produced by the instrument. The transcript contains many metaphorical descriptions of the sound. It has been described as "balls rolling in a steel pan", "water glass bell sound", or generally "pretty". Also the notion of the instrument being "on the edge" has been expressed by two participants. These reactions are also reflected in the category ratings of "Sound Quality" (5.5) and "Aesthetics" (6.3).

5 Discussion

The results show that the instrument artifact has been overall positively received. It has addressed some of the criticisms outlined in 1.1 which was also perceived by evaluators.
The usage of highly accurate and coupled physical models contributes strongly to the perceived aesthetic and control. The notion of gaining influence rather than control as described in the Post-DMI concept [17] was clearly recognized by participants who modified the coupling of the schemes.
The string provides a rich and tactile interaction method, tackling critiques of non-continuity in event-based methods (for example MIDI) [37], and lack of passive haptic feedback [10]. The results, especially those from the think-aloud protocol, confirm this through comments regarding interaction techniques and perceived acoustical vibrations.
However, claiming the full incorporation of controller and generator seems a far stretch. The need for continuous pitch tracking and shifting introduces an inherent latency (as discussed in 2.1.3). Furthermore, misclassified F0s can lead to discontinuities in the connection of the real and virtual string. This can lead to artifacts such as sudden jumps in tone height. Additionally, the pitch and general timbre of the acoustic sound produced by the string audibly deviates from the sound output of the system. This creates a further dissonance in the perception of the instrument as a whole.
The acoustical-digital hybrid nature of the instrument as such does not seem to be a sufficient condition to remutualize controller and generator. Yet, the results undeniably show that they facilitate a rich and intimate interaction with the controller. The descriptors participants used to describe the string interaction were very tangible as opposed to the notions of control and fine-tuning expressed for traditional control elements.
Acoustical-digital hybrid instruments form a conceptual framework to consider controlling and sound-producing aspects in unity during the design process. Whether or not this unity translates into the final instrument depends strongly on the execution on the concept. Constructing acoustic elements requires a deep understanding of how these elements produce sounds and precise manufacturing abilities. How the connection between acoustical and virtual is made has to be carefully tuned so that the ergotic nature of acoustic sounds is preserved while virtual elements let the musician go beyond of what is physically possible.

5.1 Future work

Future iterations of the prototype should revisit the construction of the string so that its mere acoustic sound produces a longer-lasting tone without any buzzing. This could facilitate driving the string with an actuator to create a bi-directional coupling between the acoustical and virtual elements of the instrument. Including software adaptions to dynamically drive the string, these alterations could justify a new evaluation round.
The usage of self-sensing techniques in DMIs as explored by Davison et al. [11] is also worth exploring. Compared to a string these provide more abstract affordances that could be utilized in various contexts. Additionally, they are bi-directional by definition which may help users better understand the internal state of the virtual system by acoustic and tactile means. Testing this or related techniques in a variety of smaller, more focused instrument evaluations could reveal more regarding how to make digital interactions more acoustic.
The limitations in computational performance outlined in 2.1.2 can be overcome by scaling the compute resources. Among other authors, Visi [34] demonstrated the usage of a LattePanda 3 Delta[6] which provides a strong computational foundation while staying at a small form factor. The downsides are much greater power consumption compared to the Bela, and the need for additional audio hardware in multichannel scenarios. Another option is the development of custom hardware using field programmable gate arrays (FPGAs), whose parallel capabilities are especially suited for the computation of the grid points in a FDS [15, 26].

6 Conclusion

This paper explored acoustic-digital hybrid instruments as a means to address common critiques of Digital Musical Instruments (DMIs). By integrating physical acoustic elements with physical modeling synthesis, the presented prototype demonstrates the potential to combine tangible interaction with a versatile and rich sound engine. This approach encourages exploration and creativity, offering an alternative to the limitations of purely digital systems.
However, achieving the responsiveness, intimacy, and playability of traditional acoustic instruments remains challenging. The results emphasize the importance of precise design and fine-tuning, as even minor imperfections—such as latency or mismatches between physical and digital components can disrupt the user experience.
While the current prototype shows advancements in the remutualization of the controller and generator, the results may be more akin to Laurel et. al.'s idea of separating sound from source [25], yet addressing issues of subtlety recognized by the authors. All in all, acoustic-digital hybrids present an exciting direction for future musical interface design.

[6]LattePanda3 Delta: https://www.lattepanda.com/lattepanda-3-delta

7 Ethical Standards

The project's evaluation phase involved human participants who tested the implementation. All participants provided informed consent prior to their participation. The study design and data collection adhered to ethical standards for research with human subjects. Study participation was voluntary and not remunerated.

References

[1] Jeronimo Barbosa, Joseph Malloch, Marcelo Mortensen Wanderley, and Stéphane Huot. 2015. What does" Evaluation" mean for the NIME community?. In *NIME 2015-15th International Conference on New Interfaces for Musical Expression*. Louisiana State University, 156–161. https://inria.hal.science/hal-01158080/

[2] Edgar Berdahl, Steven Backer, and Julius O. Smith III. 2005. If I had a hammer: Design and theory of an electromagnetically-prepared piano. In *ICMC*. https://www.cct.lsu.edu/~eberdahl/Papers/ICMC2005.pdf

[3] Edgar Berdahl and Julius O Smith. 2006. Some Physical Audio Effects. In *Proc. of the 9th Int. Conference on Digital Audio Effects (DAFx-06)*. Montreal, Canada.

[4] Edgar Berdahl and Julius O. Smith. 2007. Inducing Unusual Dynamics in Acoustic Musical Instruments. In *2007 IEEE International Conference on Control Applications*. 1336–1341. https://doi.org/10.1109/CCA.2007.4389421

[5] Edgar Berdahl, J. O. Smith, and Adrian Freed. 2006. Active damping of a vibrating string. *Active, Adelaide, Australia* (2006). https://ccrma.stanford.edu/~eberdahl/Papers/Active2006BerdahlSmith.pdf

[6] Stefan Bilbao. 2009. *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics* (1 ed.). Wiley. https://doi.org/10.1002/9780470749012

[7] Claude Cadoz and Marcelo Mortensen Wanderley. 2000. Gesture - Music. In *Trends in Gestural Control of Music*, Ircam-Centre Pompidou Marcelo Wanderley et Marc Battier (Ed.). https://hal.science/hal-01105543 cote interne IRCAM: Cadoz00a.

[8] Filipe Calegario. 2019. *Challenges in Designing DMIs*. Springer International Publishing, Cham, 5–17. https://doi.org/10.1007/978-3-030-02892-3_2

[9] Mads Graesboll Christensen. 2012. A method for low-delay pitch tracking and smoothing. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Kyoto, Japan, 345–348. https://doi.org/10.1109/ICASSP.2012.6287887

[10] Perry R. Cook. 2004. Remutualizing the Musical Instrument: Co-Design of Synthesis Algorithms and Controllers. *Journal of New Music Research* 33, 3 (Sept. 2004), 315–320. https://doi.org/10.1080/0929821042000317877

[11] Matthew Davison, Craig J. Webb, Michele Ducceschi, and Andrew P. McPherson. 2024. A Self-Sensing Haptic Actuator for Tactile Interaction with Physical Modelling Synthesis. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 574–581. https://nime.org/proceedings/2024/nime2024_84.pdf

[12] Alain De Cheveigné and Hideki Kawahara. 2002. YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America* 111, 4 (2002), 1917–1930.

[13] Mark Dolson. 1986. The Phase Vocoder: A Tutorial. *Computer Music Journal* 10, 4 (1986), 14–27. https://doi.org/10.2307/3680093

[14] Liam Donovan. 2018. *Travelling Wave Control of Stringed Musical Instruments*. Thesis. Queen Mary University of London. https://qmro.qmul.ac.uk/xmlui/handle/123456789/54052 Accepted: 2019-01-03T15:12:36Z.

[15] J.A. Gibbons, David Howard, and Andy Tyrrell. 2005. FPGA implementation of 1D wave equation for real-time audio synthesis. *Computers and Digital Techniques, IEE Proceedings* - 152 (Oct. 2005), 619–631. https://doi.org/10.1049/ip-cdt:20045178

[16] Gregory S. Heet. 1978. String instrument vibration initiator and sustainer. https://patents.google.com/patent/US4075921A/en

[17] Amelie Hinrichsen and Till Bovermann. 2016. Post-DMI musical Instruments. In *Proceedings of the Audio Mostly 2016*. ACM, Norrköping Sweden, 124–131. https://doi.org/10.1145/2986416.2986440

[18] T. Inagaki and J. Stahre. 2004. Human supervision and control in engineering and music: similarities, dissimilarities, and their implications. *Proc. IEEE* 92, 4 (2004), 589–600. https://doi.org/10.1109/JPROC.2004.825876

[19] Robert H. Jack, Jacob Harrison, and Andrew P. McPherson. 2020. Digital Musical Instruments as Research Products.. In *NIME*. 446–451. http://instrumentslab.org/data/jacob/DMIs_Research_Products.pdf

[20] Chris Kiefer. 2023. Dynamical complexity measurement with random projection: a metric optimised for realtime signal processing. (June 2023). https://sussex.figshare.com/articles/conference_contribution/Dynamical_complexity_measurement_with_random_projection_a_metric_optimised_for_realtime_signal_processing/23496020/2

[21] Chris Kiefer, Daniel Overholt, and Alice Eldridge. 2020. Shaping the behaviour of feedback instruments with complexity-controlled gain dynamics. In *Proceedings of the New Interfaces for Musical Expression 2020 conference (NIME Proceedings)*. International Conference on New Interfaces for Musical Expression, 343–348. https://nime2020.bcu.ac.uk/ 20th International Conference on New Interfaces for Musical Expression, Nime 2020 ; Conference date: 21-07-2020 Through 25-12-2020.

[22] Annie Luciani, Jean-Loup Florens, Damien Couroussé, and Julien Castet. 2009. Ergotic Sounds: A New Way to Improve Playability, Believability and Presence of Virtual Musical Instruments. *Journal of New Music Research* 38, 3 (Sept. 2009), 309–323. https://doi.org/10.1080/09298210903359187

[23] Thor Magnusson and Enrike Hurtado Mendieta. 2007. The acoustic, the digital and the body: a survey on musical instruments. In *Proceedings of the 7th international conference on New interfaces for musical expression - NIME '07*. ACM Press, New York, New York, 94. https://doi.org/10.1145/1279740.1279757

[24] Romain Michon, Julius Orion Smith, Matthew Wright, Chris Chafe, John Granzow, and Ge Wang. 2017. Mobile Music, Sensors, Physical Modeling, and Digital Fabrication: Articulating the Augmented Mobile Instrument. *Applied Sciences* 7, 1212 (Dec. 2017), 1311. https://doi.org/10.3390/app7121311

[25] Laurel Smith Pardue, Kurijn Buys, Michael Edinger, Daniel Overholt, and Andrew McPherson. 2019. Separating sound from source: sonic transformation of the violin through electrodynamic pickups and acoustic actuation. In *NIME 2019 New Interfaces for Musical Expression conference*. 278–283.

[26] Florian Pfeifle and Rolf Bader. 2012. Real-time finite difference physical models of musical instruments on a field programmable gate array (FPGA). In *Proc. of the 15th Int. Conference on Digital Audio Effects (DAFx-12)*. 17–21. https://dafx.de/paper-archive/2012/papers/dafx12_submission_47.pdf

[27] P. J. Charles Reimer and Marcelo M. Wanderley. 2021. Embracing Less Common Evaluation Strategies for Studying User Experience in NIME. In *International Conference on New Interfaces for Musical Expression*. https://doi.org/10.21428/92fbeb44.807a000f

[28] Joshua D. Reiss and Andrew P. McPherson. 2015. *Audio Effects: Theory, Implementation and Application*. CRC Press, Boca Raton, Fla.

[29] Daniel Schlessinger and Julius O. Smith. 2009. The Kalichord : A Physically Modeled Electro-Acoustic Plucked String Instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Pittsburgh, PA, United States, 98–101. https://doi.org/10.5281/zenodo.1177671

[30] Gian-Marco Schmid. 2015. *Evaluation of musical instruments from the musician's perspective*. PhD Thesis. University of Basel Basel, Switzerland. http://maschiiine.com/MA/2015_Schmid.pdf

[31] Julius O. Smith. 2004. Virtual Acoustic Musical Instruments: Review and Update. *Journal of New Music Research* 33, 3 (Sept. 2004), 283–304. https://doi.org/10.1080/0929821042000317859

[32] D. Stowell, A. Robertson, N. Bryan-Kinns, and M. D. Plumbley. 2009. Evaluation of live human–computer music-making: Quantitative and qualitative approaches. *International Journal of Human-Computer Studies* 67, 11 (Nov. 2009), 960–975. https://doi.org/10.1016/j.ijhcs.2009.05.007

[33] Halldór Úlfarsson. 2018. The halldorophone: The ongoing innovation of a cello-like drone instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 269–274.

[34] Federico Visi. 2024. The Sophtar: a networkable feedback string instrument with embedded machine learning. In *International Conference on New Interfaces for Musical Expression (NIME), Utrecht, Netherlands, September 4-6, 2024*. International Conference on New Interfaces for Musical Expression, 142–148. https://www.diva-portal.org/smash/record.jsf?pid=diva2:1916497

[35] Vesa Välimäki and Tapio Takala. 1996. Virtual musical instruments—natural sound using physical models. *Organised Sound* 1, 2 (1996), 75–86.

[36] Marcelo Mortensen Wanderley and Nicola Orio. 2002. Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal* 26, 3 (2002), 62–76.

[37] David Wessel and Matthew Wright. 2002. Problems and prospects for intimate musical control of computers. *Computer music journal* 26, 3 (2002), 11–22.

[38] Silvin Willemsen, Stefan Bilbao, Michele Ducceschi, and Stefania Serafin. 2021. Dynamic Grids for Finite-Difference Schemes in Musical Instrument Simulations. In *2021 24th International Conference on Digital Audio Effects (DAFx)*. 144–151. https://doi.org/10.23919/DAFx51585.2021.9768302

[39] Silvin Willemsen and Stefania Serafin. 2022. Real-time implementation of the dynamic stiff string using finite-difference time-domain methods and the dynamic grid. In *Proceedings of the 25th International Conference on Digital Audio Effects (DAFx20in22),, G. Evangelista and N. Holighaus, Eds.* 130–137. https://www.researchgate.net/profile/Silvin-Willemsen/publication/363641614_Real-Time_Implementation_of_the_Dynamic_Stiff_String_using_Finite-Difference_Time-Domain_Methods_and_the_Dynamic_Grid/links/6326ffa7873eca0c0098cd4e/Real-Time-Implementation-of-the-Dynamic-Stiff-String-using-Finite-Difference-Time-Domain-Methods-and-the-Dynamic-Grid.pdf

A Online ressources

A.1 Demonstrational video

A demonstration of the second artifact can be found at: https://vimeo.com/1053084541/550b428fdd

A.2 Source code

The source code and other relevant design files can be found at: https://github.com/moewe-audio/NEMO-acoustic-digital-hybrid-instrument