Temporal Synchronization of Multimodal Hyperscanning Recordings: Challenges, Methodologies, and Best Practices

Simon L. Kappel[1], Andreas Nørtoft Jørgensen[1] and Preben Kidmose[1]

***Abstract*— In multimodal hyperscanning, multiple signal modalities are collected simultaneously from multiple participants. Temporal synchronization of sensor nodes in such scenarios is crucial for analyzing temporal dependencies across participants and modalities.**

Here, we present an overview of challenges, methods and best practices for achieving precise temporal synchronization in multimodal hyperscanning studies. We discuss various methods for temporal synchronization, the challenges involved, and the implications for data analysis and interpretation.

We also present analysis and considerations on the effects of sampling clock mismatches between nodes in a system, and provide guidance on how to quantify trigger offsets and noise in a recording setup.

By providing an overview of different methods, we aim to guide researchers in optimizing their experimental designs for successful multimodal hyperscanning studies.

I. Introduction

Temporal synchronization of data is crucial in hyperscanning, where brain activity from multiple participants are recorded simultaneously [1]. Analyses of hyperscanning data often rely on inter-subject comparisons, with temporal analysis methods such as correlation-based techniques commonly employed to investigate inter-brain connectivity and synchronization [2], [3]. Also, methods based on phase and envelope coupling are frequently employed to assess the temporal relationships between neural signals from different participants [4], [5].

Moreover, when conducting recordings with a multimodal setup, temporal alignment between modalities is vital to obtain a comprehensive understanding of how different mechanisms, such as physiological and behavioral processes, are temporally connected [6]. This is particularly important in multimodal hyperscanning setups, where data from various sources, such as EEG, fNIRS, pulse oximetry, and behavioral measures, need to be synchronized accurately both within the participant and between participants, to provide a holistic view of the interactions between different systems.

The goal is often to achieve sample-level synchronization [7], [8], where recordings are synchronized in time, sample-by-sample. However, in some cases, it might be more relevant to align the data to the time where the subject receive the stimulus of interest. For example, if the stimulus is an auditory signal presented on a stage, the distance to the stage will influence when the stimulus reaches the participant. In such scenarios, it might be more pertinent to align data in a hyperscanning study to the time where the subject receive the stimulus rather than the time of its presentation.

In summary, careful consideration of temporal synchronization is paramount in hyperscanning research to ensure the validity and reliability of the findings. By addressing these challenges, researchers can better understand the temporal dynamics of inter-brain connectivity and the interplay between different physiological and behavioral processes.

In this paper we will focus on the methods and techniques for achieving precise temporal synchronization in hyperscanning studies, with a particular emphasis on multimodal setups. We will discuss the different methods for temporal synchronization, the challenges involved, and the implications for data analysis and interpretation. By providing an overview of the different methods, we aim to guide researchers in optimizing their experimental designs for conducting successful Hyperscanning studies.

II. Methods

A. Temporal synchronization

When analysis is based on temporal comparison or averaging of features from multiple data sources, it is important to ensure that the sources are properly synchronized to obtain meaningful results. Depending on the features of interest, different types of synchronization might be preferred. Two primary methods of synchronization are relevant to highlight here: sample-based synchronization and stimulus-based synchronization.

In some cases, sample-based and stimulus-based synchronization might be interchangeable. However, in other experimental setups, the propagation time of stimuli or behavioral signals may not be negligible. For instance, when the stimulus is an acoustic signal, differences in the propagation path length from the sound source to individual participants will result in different delays, as illustrated in Fig. 1. Another example is scenarios involving interacting participants, where the interaction is based on behavior communicated over a communication channel, where delays in the communication channel may not be negligible. Such delays can impact not only the synchronization of data but also the interaction dynamics themselves.

Considering the setup in Fig. 1, the time of presentation and perception of visual stimuli would be almost identical. However, consider auditory stimulation and assume that $l_1 = 10\,\mathrm{m}$ and $l_2 = 1\,\mathrm{m}$. Given a sound propagation speed of $v = 343\,\mathrm{m/s}$, the difference in time of perception for the two participants will be $\Delta t = (l_1 - l_2)/v = 26\,\mathrm{ms}$. Such a

*This research was supported by the Center for Ear-EEG, Department of Electrical and Computer Engineering, Aarhus University.

[1]S. L. Kappel, A. N. Jørgensen and P. Kidmose are with Department of Electrical and Computer Engineering, Aarhus University, DK-8200 Aarhus N, Denmark.

Manuscript 764 submitted to 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) . Received February 9, 2025.

time difference would be significant in many physiological responses and should be considered when aligning data in such a setup.

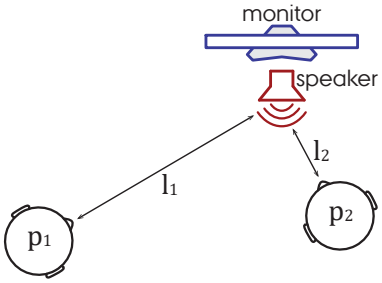


Fig. 1: Spatial setup with two participants, p_1 and p_2, listening to an audio source (speaker) and viewing a visual source (monitor).

B. Methods for sample-level synchronization

Synchronization of data can be performed using various practical implementations. However, the common goal is to periodically realign data to compensate for potential mismatches in the sampling performed by the different sensor nodes. Here, we highlight four methods to obtain sample-level synchronization. The methods are illustrated in Fig. 2(a) to (d).

(a) **Clock based:** In this method, the sensor nodes share a common sampling clock, ensuring synchronization on a sample-by-sample basis. This approach provides the highest precision of the temporal synchronization, but is rarely supported by commercial equipment.

(b) **Trigger based:** The most common method is to use stimulus locked events (i.e. triggers). In this method, the same trigger signal is transmitted to all sensor nodes and sampled together with the sensor data. Synchronization of data can then be performed at sample-level when a trigger appears in the recorded data.
Typically, the trigger signal is transmitted via cables. However, in larger setups, such as large-scale Hyperscanning, cable-based alignment can be cumbersome to set up and may be a possible path for noise interference. In such cases, wireless trigger distribution might be more feasible, using methods like infrared light or low-latency RF protocols such as Zigbee.

(c) **RTC based:** Synchronization based on a real-time clock (RTC) involves equipping each sensor node with an RTC to timestamp the data. The timestamps are then used to align the data during post-processing. The accuracy of this method depends on clock mismatches between and synchronization of the RTCs across sensor nodes. To obtain the best synchronization, nodes should be synchronized to a common time source just before the experiment. This method does not require communication to the sensor nodes during the experiment, making it particularly advantageous for larger setups.

(d) **LAN based:** Synchronization based on a local area network (LAN) can be achieved using various methods, typically relying on the network clock for synchronization. One of the most widely used methods is the Lab Streaming Layer (LSL), developed by Kothe et al. [9]. With well-characterized network and acquisition equipment, LSL can achieve near sample-level synchronization [9], [10].

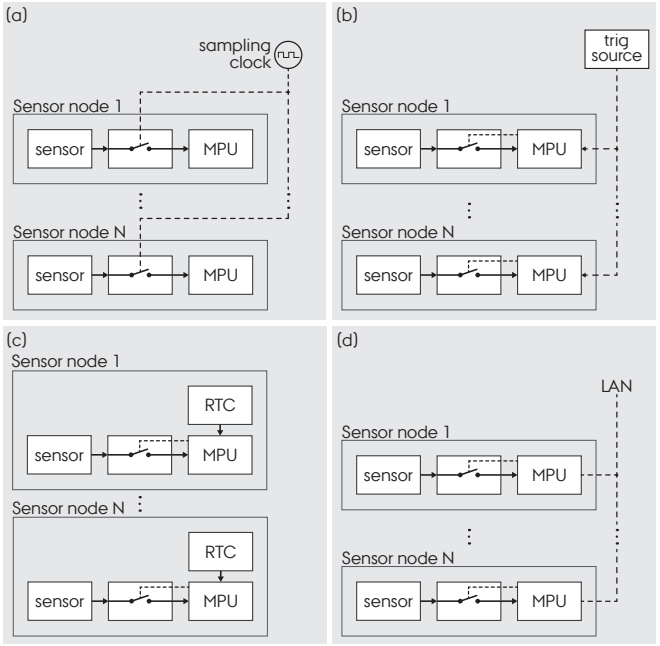


Fig. 2: Illustration of sample-level synchronization methods. Each sensor node samples data from a sensor, for processing in the main processing unit (MPU) of the sensor node, (a) Clock based synchronization, (b) Trigger based synchronization, (c) RTC based synchronization, and (d) LAN based synchronization.

C. Trigger encoding

In most settings, it is advantageous to send an event ID with the trigger signal. When multiple trigger inputs are available on the sensor node, the ID can be encoded as a parallel signal. However, when only a single trigger input is available, serial encoding of the event ID can be utilized. Manchester coding (or similar methods) can be employed to make the transmission robust to hardware errors and to enable a variable word length (number of bits) in the event ID. Fig. 3 illustrates Manchester encoding, where the event

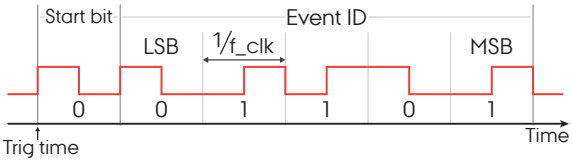


Fig. 3: Illustration of Manchester coding used for robust trigger signal transmission.

Manuscript 764 submitted to 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) . Received February 9, 2025.

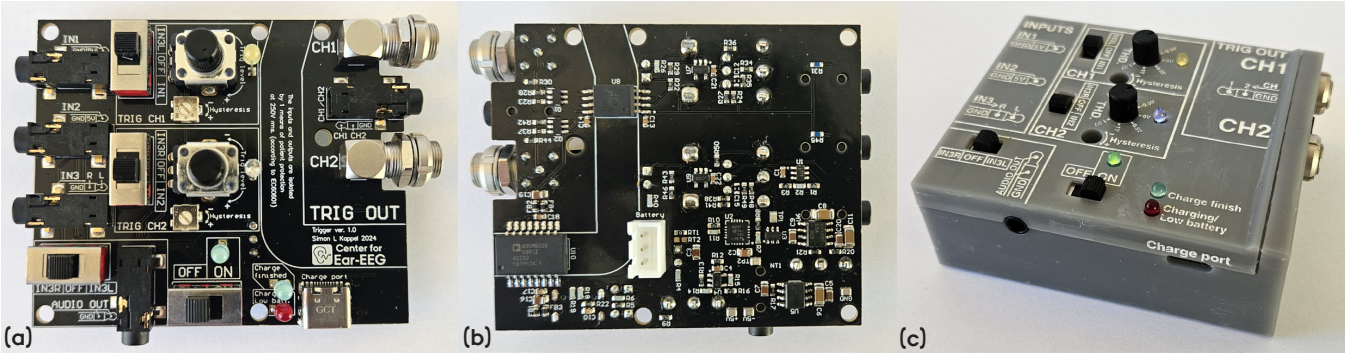


Fig. 4: Implementation of the designed trigger box (a) PCB top side, (b) PCB bottom side, (c) PCB mounted in a 3D printed casing together with a Lithium Ion battery. Design files for the trigger box is available from ece.au.dk/triggerbox [11].

ID is encoded with 5 bits, and the rising edge of the start bit indicates the time of the trigger. Manchester encoding is based on a clock, with a frequency, f_clk, which should be below half of the sampling rate of the sensor node with the lowest sampling rate, to ensure correct decoding of the trigger. With Manchester coding, each bit period always has a transition in the middle, and therefore the end of the event ID is represented by more than one clock period without a transition. For a Matlab-based implementation of Manchester encoding/decoding, see gitlab.au.dk/tech_ear-eeg/serialtrigcoder_matlab.

D. Trigger sources

Triggers can be generated from various sources, typically depending on the experimental paradigm. Triggers often indicate events that are temporally synchronized with the stimulus of interest for data analysis. Common trigger sources include:

- Digital, e.g. from a computer presentings the stimulus.
- Sound card, which enable synchronization between audiotory stimulus and triggers.
- Optical sensors, which can detect light changes, e.g. from a monitor.
- Buttons, including mechanical, resistive (FSR), and capacitive buttons.

E. Trigger interference

Triggers are often aligned with the response of interest, and therefore, trigger noise in the recorded physiological signal can be mistaken for a physiological response that is not actually present. Thus, it is crucial to avoid trigger-related interference in the physiological signal. Several measures can be taken to minimize trigger noise in the recorded signal, including:

- Galvanic isolation between the trigger source and the sensor nodes.
- Considerate cable routing to optimize spatial distance between trigger cables and sensor cables, especially avoiding trigger and sensor cables to run in parallel.
- Battery powering the trigger source to avoid ground loops.
- Using shielded or twisted pair cables for trigger signals.

Additionally, when working with steady-state responses, the trigger frequency can be decreased to reduce the power of potential trigger noise at the response frequency. Furthermore, the temporal alignment between the response and the triggers can be eliminated by using an aperiodic trig signal or randomizing the trigger times. The timing of the stimulus can be determined retrospectively using the known structure of the trigger times.

F. Trigger box

Generally, the different trigger sources must be converted to a binary signal to be parsed by most commercial amplifiers. This can be achieved by using thresholding (preferably with hysteresis) of the analog signal, to obtain a digital signal that can be fed to the sensor node. This is typically done with a so-called trigger box. Preferred features of a trigger box include:

- Galvanic isolation between input and output to avoid ground loops and provide electrical safety.
- Battery powered, to minimize power line interference to the trigger signal.

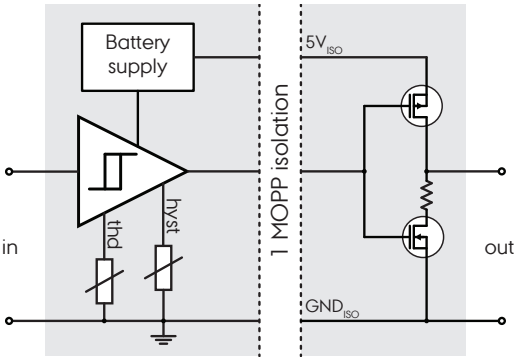


Fig. 5: Input-to-output function of the designed trigger box. The input goes to a comprator, which supports threshold (thd) and hysteresis (hyst) adjustments. The output from the comperator is directed to a digital isolator, connected to a push-pull output. In addition, the supply from the input stage is galvanically isolated from the output stage, as illustrated with the $5V_{ISO}$ and GND_{ISO} supply.

- Low-latency from trigger input to trigger output.
- Adjustable threshold and hysteresis.

To meet these requirements, we designed at trigger box with the input-to-output function shown in Fig 5. The input signal is fed to a comperator, for which the threshold and hysteresis can be adjusted with two potentiometers. The output from the comperator is passed through a digital isolator, complying with 1 means of patient protection (MOPP), according to the IEC60601-1 standard. In addition, the supply from the input stage is also galvanically isolated from the output stage with 1 MOPP. The output features a push-pull stage, with a maximum output current of 50mA. This functionality was implemented on a printed circuit board (PCB), containing two identical channels, each corresponding to Fig. 5, along with related connectors and switches, to obtain flexible input and output connections. The final PCB is illustrated in Fig. 4(a) and (b) along with a picture of the PCB mounted in a 3D printed casing in Fig. 4(c). The design files for the trigger box is available from ece.au.dk/triggerbox [11].

G. Effect of sample clock mismatch

Mismatch in sample clock between sensors nodes will always be present, unless the sensor nodes share the sampling clock. Mismatches can be minimized by using clock sources with low tolerance and drift. As described in Section II-B, triggers can be utilized to realign data in the post-processing. However, if the interval between triggers is too large, misalignment of data between triggers might still be an issue. Synchronization of recordings with RTC time stamps corresponds to having only one trigger in the beginning of the recording to support the synchronization. Thus, mismatch in sampling clocks between sensor nodes synchronized with RTC will developed over the course of the recording.

The effect of mismatch between the sampling clock of two sensor nodes can be expressed as the correlation between recordings from the two sensor nodes. As an example, lets consider two sinusoidal signals, $S_1(n)$ and $S_2(n)$, sampled with sampling clocks F_{S1} and F_{S2}, respectively. The correlation coefficient between the two signals can be expressed as

$$r_{12}(0) = \frac{1}{N}\sum_{n=0}^{N} S_1(n)\cdot S_2(n) \tag{1}$$

where, $S_1(n) = \sin(\omega \cdot n/F_{S1})$, $S_2(n) = \sin(\omega \cdot n/F_{S2})$, $F_{S2} = F_{S1}\cdot(1 + a\cdot 10^{-6})$, and ω is the angular frequency of the sinusoid.

Inserting the expression for $S_1(n)$ and $S_2(n)$, results in

$$\begin{aligned} r_{12}(0) =& \frac{1}{N}\cdot\frac{1}{2}\sum_{n=0}^{N}\cos\left(\omega\cdot n\cdot\left(\tfrac{1}{F_{S1}}-\tfrac{1}{F_{S2}}\right)\right) \\ &-\cos\left(\omega\cdot n\cdot\left(\tfrac{1}{F_{S1}}+\tfrac{1}{F_{S2}}\right)\right) \end{aligned} \tag{2}$$

For $N \gg \omega/(2\pi)$, the last term in equation 2 is neglectable, and the correlation reduces to

$$r_{12}(0) = \frac{1}{2N}\sum_{n=0}^{N}\cos\left(\omega\cdot n\cdot\left(\tfrac{1}{F_{S1}}-\tfrac{1}{F_{S2}}\right)\right) \tag{3}$$

If we consider this as a Riemann sum approximation of the integral of a cosine, equation 3 can be expressed as

$$r_{12}(0) = \frac{1}{2N}\cdot\frac{1}{\omega\left(\frac{1}{F_{S1}}-\frac{1}{F_{S2}}\right)}\cdot\sin\left(\omega\cdot n\cdot\left(\tfrac{1}{F_{S1}}-\tfrac{1}{F_{S2}}\right)\right) \tag{4}$$

The correlation at $F_{S2} = F_{S1}$ is $r_{11}(0) = \frac{1}{2}$. Thus, the normalized correlation can be expressed as

$$\hat{r}_{12}(0) = \left|\frac{r_{12}(0)}{r_{11}(0)}\right| = \left|\frac{\sin\left(\omega\cdot n\cdot\left(\tfrac{1}{F_{S1}}-\tfrac{1}{F_{S2}}\right)\right)}{\omega\cdot N\cdot\left(\tfrac{1}{F_{S1}}-\tfrac{1}{F_{S2}}\right)}\right| \tag{5}$$

From equation 5 we see that the normalized correlation depends on

1) The mismatch in the sampling clocks, $(1/F_{S1} - 1/F_{S2})$
2) The duration of the signal, N
3) The frequency of the sinusoidal signals, ω.

III. Results and Discussion

In the following, we elaborate on the effects of sample clock mismatch, methods to quantify trigger offsets, and trigger interence in the recording setup.

A. Effects of sample clock mismatch

Fig. 6 illustrates the effect of sample clock mismatch between sensor nodes at a frequency of 20 Hz. The figure shows the normalized correlation, $\hat{r}_{12}(0)$, as a function of clock mismatch (in ppm) for various analysis windows. It demonstrates how a mismatch has a larger effect when the window is increased, as the mismatch accumulates over time.

Fig. 7 illustrates the effect of frequency for a clock mismatch of 100 ppm between sensor nodes. The figure shows the normalized correlation, $\hat{r}_{12}(0)$, as a function of the frequency for various analysis windows. It illustrates how the effect of clock mismatch is more pronounced for larger analysis windows and higher frequencies. Correlation-based measures are commonly used in hyperscanning studies to investigate inter-brain connectivity and synchronization [2]–[4]. Figs. 6 and 7 highlight the importance of precise

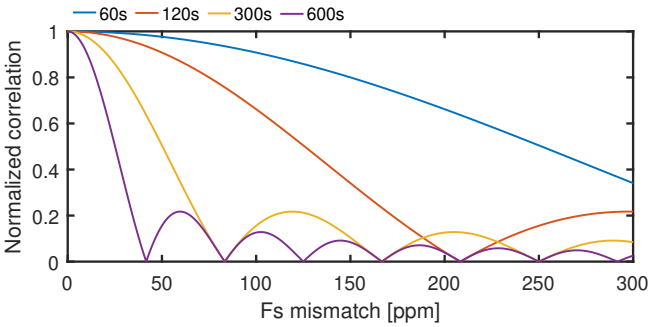


Fig. 6: Effect of clock mismatch between nodes, at a frequency of 20 Hz. The normalized correlation, $\hat{r}_{12}(0)$, is shown as a function of sampling clock mismatch (in ppm) for various analysis windows, as given in the legend.

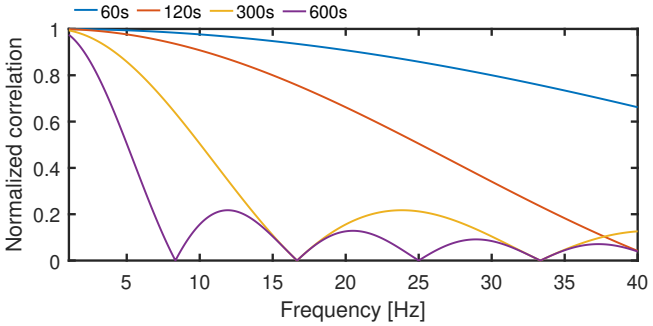
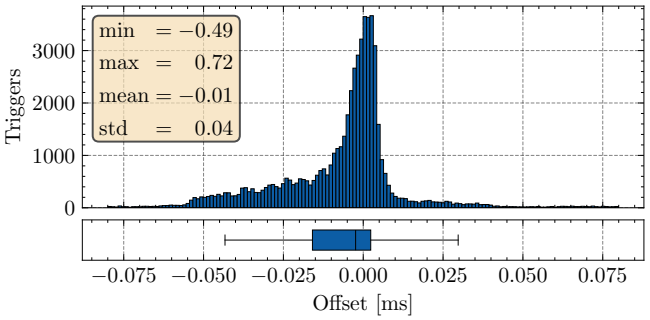


Fig. 7: Effect of frequency, for a clock mismatch of 100ppm between sensor nodes. The normalized correlation, $\hat{r}_{12}(0)$, is shown as a function of the frequency for various analysis windows, as given in the legend.

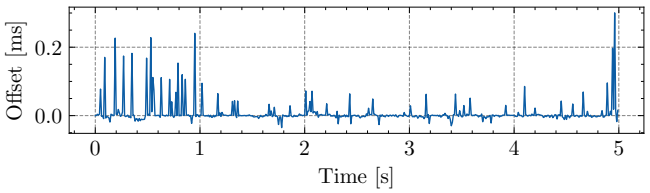
synchronization between sensor nodes to obtain valid and reliable results in hyperscanning studies. The results show that the effect of clock mismatch is more pronounced for larger analysis windows and higher frequencies. Therefore, it is crucial to carefully consider the needed synchronization strategy for a study, to ensure the necessary synchronization within the analysis window of interest. For example, if the required sampling clock mismatch cannot be achieved, trigger or LAN-based synchronization can be utilized, as described in Section II-B. If the necessary synchronization is not obtained, the response amplitude might be reduced as a consequence, and thus may not reflect the true physiological response amplitude.

B. Quantifying trigger offsets

When working with RTC or LAN-based synchronization of sensor nodes, unexpected and varying offset latencies can occur. Therefore, it is crucial to characterize the setup before the experiment to ensure that the synchronization meets the requirements. One approach is to use a cable-based trigger as a reference point for the characterization. Here, we generated a trigger at 100 Hz using a waveform generator (Agilent 33500B Series) connected to two sensor nodes (Raspberry Pi Zero 2 W), which were timestamping and streaming data via LSL over a wireless LAN. To avoid including clock mismatches between sensor nodes in the characterization, the sensor nodes shared the same sampling clock. The trigger signal was initiated right after starting the recording and was deactivated just before the recording ended. The offsets between sensor nodes were then calculated trigger by trigger. Fig. 8(a) shows a histogram of the calculated offsets between two nodes from a ten minutes long recording, and in Fig. 8(b) is plotted five seconds of the offsets. Both sensor nodes in the experiment were sampling at 500 Hz, and thus, the variation in the offsets is well within +/- 1 sample period, indicating sample-level synchronization. To ensure representative data, it is recommended to perform such characterization of offsets for all sessions of an experiment and with the full experimental setup, including recording devices, the network, and the computers streaming and receiving LSL streams.



(a) Histogram of the offsets between two nodes, from a ten minutes long recording.



(b) Five seconds of the offsets between two nodes.

Fig. 8: Characterization of trigger offsets between two sensor nodes.

C. Characterization of trigger interference in the recording setup

Often, triggers are phase-locked to the signal of interest, and therefore trigger-related artifacts in the data can be difficult to distinguish from the response of interest. Thus, it can be important to perform a negative control to characterize the level of trigger noise in a recording setup and to rule out issues with the analysis pipeline. A common practice is to perform a recording that represents the final setup, but without presenting the stimulus to the participant. The data can then be analyzed with the analysis pipeline, to confirm the absence of false reponses due to trigger interference.

IV. Conclusion

In this paper, we have discussed the importance of precise temporal synchronization in hyperscanning studies, with a particular emphasis on multimodal setups. The paper has provided an overview of the different methods for achieving temporal synchronization, the challenges involved, and the implications for data analysis and interpretation. Addressing these aspects is anticipated to contribute to the improvement of experimental research in multimodal and hyperscanning studies, thereby advancing the understanding of temporal dynamics of inter-brain connectivity and the interplay between physiological and behavioral processes.

References

[1] P. R. Montague, G. S. Berns, J. D. Cohen, S. M. McClure, G. Pagnoni, M. Dhamala, and M. C. Wiest, "Hyperscanning: Simultaneous fMRI during linked social interactions," *NeuroImage*, vol. 16, no. 4, pp. 1159–1164, 2002.

Manuscript 764 submitted to 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) . Received February 9, 2025.

[2] U. Hakim, S. De Felice, P. Pinti, X. Zhang, J. Noah, Y. Ono, P. W. Burgess, A. Hamilton, J. Hirsch, and I. Tachtsidis, "Quantification of inter-brain coupling: A review of current methods used in haemodynamic and electrophysiological hyperscanning studies," *NeuroImage*, vol. 280, p. 120354, Oct. 2023.
[3] G. Dumas, F. Lachat, J. Martinerie, J. Nadel, and N. George, "From social behaviour to brain synchronization: Review and perspectives in hyperscanning," *IRBM*, vol. 32, no. 1, pp. 48–53, 2011.
[4] A. Zamm, J. D. Loehr, C. Vesper, I. Konvalinka, S. L. Kappel, O. A. Heggli, P. Vuust, and P. E. Keller, "A practical guide to EEG hyperscanning in joint action research: From motivation to implementation," *Social Cognitive and Affective Neuroscience*, vol. 19, no. 1, p. nsae026, May 2024.
[5] A. Ayrolles, F. Brun, P. Chen, A. Djalovski, Y. Beauxis, R. Delorme, T. Bourgeron, S. Dikker, and G. Dumas, "HyPyP: A Hyperscanning Python Pipeline for inter-brain connectivity analysis," *Social Cognitive and Affective Neuroscience*, vol. 16, no. 1-2, pp. 72–83, Jan. 2021.
[6] S. Makeig, K. Gramann, T.-P. Jung, T. J. Sejnowski, and H. Poizner, "Linking brain, mind and behavior," *International Journal of Psychophysiology*, vol. 73, no. 2, pp. 95–100, Aug. 2009.
[7] P. Barraza, G. Dumas, H. Liu, G. Blanco-Gomez, M. I. van den Heuvel, M. Baart, and A. Pérez, "Implementing EEG hyperscanning setups," *MethodsX*, vol. 6, pp. 428–436, Feb. 2019.
[8] F. Babiloni and L. Astolfi, "Social neuroscience and hyperscanning techniques: Past, present and future," *Neuroscience & Biobehavioral Reviews*, vol. 44, pp. 76–93, Jul. 2014.
[9] C. Kothe, S. Y. Shirazi, T. Stenner, D. Medine, C. Boulay, M. I. Grivich, T. Mullen, A. Delorme, and S. Makeig, "The lab streaming layer for synchronized multimodal recording," *bioRxiv*, Feb. 2024.
[10] S. Dasenbrock, S. Blum, P. Maanen, S. Debener, V. Hohmann, and H. Kayser, "Synchronization of ear-EEG and audio streams in a portable research hearing device," *Frontiers in Neuroscience*, vol. 16, p. 904003, Sep. 2022.
[11] S. L. Kappel, A. Nørtolt Jørgensen, and P. Kidmose, "Triggerbox design files," https://ece.au.dk/triggerbox.