nature portfolio

Peer Review File

Secure wireless communication of brain-computer interface and mind control of smart devices enabled by space-time-coding metasurface

Corresponding Author: Professor Tie Jun Cui

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The proposed work involves a brain-machine interface combined with space-time coding metasurfaces for wireless communication. However, the core novelty of the research is unclear, as it primarily seems to focus on combining EEG data with metasurfaces. Additionally, the paper does not address security measures for the wireless communication aspect, and providing quantitative metrics, such as bit error rate (BER) for the wireless links, would be beneficial. Moreover, there are several typos in the article, so it is recommended that the paper be thoroughly proofread.

Reviewer #2

(Remarks to the Author)
This work integrates brain-computer interfaces with spatiotemporal metasurfaces, which is quite innovative. A harmonic encryption method is employed to achieve secure wireless communications. Although the combination of brain-computer interfaces and metasurfaces has been explored for several years, most of the researches focused on static metasurfaces. The application of spatiotemporal metasurfaces introduces various harmonics, which can be utilized for encryption and wireless communications, as demonstrated in the main text. The experiments presented are compelling and impressive. Overall, this work is fancy and merits acceptance. There are some optional suggestions to enhance the manuscript:

The introduction section contains a considerable amount of detail, particularly in the last paragraph. For clarity, the authors are kindly suggested that the authors condense this section.

Reviewer #3

(Remarks to the Author)
The article is very difficult to read. It took me an entire afternoon to go through it, paragraph by paragraph, while also referring to the supplemental materials. Despite this effort, many aspects remain unclear to me. I understand the overall purpose of the work and some individual steps, but I do not have a comprehensive understanding of the process involved in achieving secure wireless communications.

The authors attempt to combine several popular techniques, such as reconfigurable intelligent surfaces, machine learning (ML), and the Metaverse. While I am not suggesting that this approach is without merit, it is very confusing, especially for readers who are not experts in secure wireless communications.

One of the biggest difficulties is the lack of focus on key features. Besides the very general overview provided in Figure 1, which is not self-explanatory, the article does not emphasize the crucial points. Instead, all arguments are presented at the same level, leaving the reader to discern the most important aspects. The authors seem to expect the reader to do the work of prioritizing the content.

The article also uses many acronyms (for example, STC is introduced in the abstract but only partially defined), which further complicates understanding.

Moreover, several concepts are mixed together in a way that adds to the confusion. For instance, the use of visual excitation for encoding information is not well explained. It is unclear how this process works—does the person in front of the LED panel select different panels in sequence? Or is it merely a method to monitor which part of the panel is being observed? Additionally, there seem to be as many LEDs as unit cells on the metasurface, yet these two effects appear to be completely decoupled. Why is space-time coding using a metasurface chosen? Why is the information transmitted to two users? Is the same information transmitted to both users? If so, why two users and not one or more?

The article also mentions that the pixel is split in two, which seems impossible. Instead of this complex setup, why not simply use a classical antenna or an array of antennas? Why are transmissions made over several harmonics? Isn't the focusing effect sufficient?

The discussion about the keys is also unclear—are they known to the receivers?

The second part of the article, which discusses control, is equally confusing. Is there any security feature involved, or is it merely a demonstration of remote control using EEG?

In conclusion, it is very difficult to judge the effectiveness of the work because it is so confusing and mixes so many different concepts without providing the necessary keys to understand not only the choices made but also how the system works. For these reasons, I recommend rejecting the paper.

Specific Comments:

- Line 77: "Hence, the STC metasurface is a … ." I do not understand the sentence. Is the STC processing the information? How?
- Line 84: "Information interaction" – What does this mean?
- Line 86: "High-security" – How high? Is it equivalent to 256-bit AES encryption? (I don't think so.)
- Line 100: "Regulate the EM" – EM produced by what?
- Line 110: "The BSTMCM system empowers individuals." What does this mean?
- Line 111: "Abundant harmonic…" – This needs clarification.
- Line 115: "Transmitted information" – But which information? Sent by the user, or by the BSTCM?
- Line 116: "Visual secret ciphertexts" – Are these chosen by the user or by the BSTCM? If the latter, what is the interest?
- Line 122: I do not understand the Metaverse applications.
- Line 124: "There are now secret keys?" Are these different from the ciphertexts?
- Line 131-132: There is repetition here.
- Line 133: The spectrum is not completely clear. What is the difference between notes 2 and 5—recognition or extraction?
- Line 173: "Such results underscore the efficacy of the SSVEP component…." I do not understand—was ML used or not?
- Line 186: Why are the LEDs connected to the PIN diodes? Would it not have been simpler to decouple the two?
- Line 246: The article speaks of images without explaining that images are being transmitted.
- Lines 254-305: I do not understand where the brain coding occurs.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
While the revision has addressed some of the previous questions, it remains unclear how information is transmitted from the EEG signal to other users.

It appears that all information transfer is initiated by the LED lights. Does the host user have the ability to freely transmit data of their choice? If not, what is the justification for using a Brain-Machine Interface (BMI)? In this case, wouldn't a standard ST metasurface multi-user scheme be sufficient?

Additionally, the data transmission process needs further clarity, particularly regarding the amount of data that can be transmitted in terms of bits. Furthermore, a BER of 30% is quite high for the legal user (Fig.R2), which raises concerns about the system's reliability and practical usability.

Reviewer #3

(Remarks to the Author)
The writing of the resubmitted article is much easier to understand, at least up to the section on secure communication. However, this part remains highly confusing, and the provided corrections and accompanying letter do not sufficiently clarify the content. Once again, I spent a significant amount of time—nearly an entire day—trying to understand the coding process

described. Several critical issues persist:
1. Visual Secret Sharing (VSS):
◦ Why is this termed "visual"? Beyond the secure transmission section, the role of visual stimuli is not explained or justified.
2. Transmission of VSK1 and VSK2:
◦ To my understanding, VSK1 and VSK2 are transmitted to user 1 and user 2 using the +1 and -1 harmonics, respectively, as per the coding schemes illustrated in Figures 4j to 4m. However, this is not explicitly stated in the text.
3. Harmonics:
◦ The frequency f0=1kHz is much smaller than the coherence frequency of classical propagation environments. Therefore when the STC metasurface beams in a specific direction with a commuting frequency fc, the signal levels at fc+f0 and fc–f0 are the same (refer to Eq. 27 in the supplemental material). This raises the question: what is the utility of distinguishing between the +1 and -1 orders? Is it not sufficient to distinguich between focusning on user 1 and 2.
4. Generation and Role of HSK1 and HSK2:
◦ How are HSK1 and HSK2 generated? How are they used for decryption? Are VSK1 and VSK2 sufficient on their own? These points are unclear.
5. Transmission of K1 and K2:
◦ According to Figure 25 in the supplemental material, transmitting K1 and K2 appears to require a separate transmission system. If so, why not use this system for transmitting the primary information?
6. User Communication for Decryption:
◦ Do users need to communicate with each other to decrypt the signal? If so, how is this communication established and managed?
7. Encoding/Decoding Process:
◦ The method for encoding and decoding information using K1 and K2 remains vague and inadequately explained.
Conclusion
While I find the results concerning the control of smart devices compelling, the same cannot be said for the secure wireless communication aspects. In addition to an incomplete and unclear description of the transmission process, there are fundamental issues that render the approach questionable:
• The rationale for using encoding based on visual stimuli (as claimed in some parts of the manuscript) is unclear and appears unnecessary. A robust random generator would likely achieve better efficiency and simplicity.
• The overall communication system appears overcomplicated, seemingly incorporating as many techniques as possible (e.g., ML, FPGA, STC matrix, secure communication, IoT, etc.), which results in a lack of consistency and coherence.
For these reasons, I still must recommend rejecting the paper. While the communication approach shows some promise, it requires significant revision, focusing on removing unnecessary steps and providing a clearer, more streamlined methodology.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The revision has addressed the reviewer previous comments.

Reviewer #3

(Remarks to the Author)
The new version of the manuscript is indeed clearer and easier to work.
However, several new typos have been introduced, which gives the impression that even the authors might be somewhat lost in their own scientific framework.
• For example, line 266: the sequence “M2-M2-M0-M0-M2-M2-M0-M0” is repeated twice without clear justification.
• Line 269: the reference should be to Figure 4, not Figure 3.
• Line 257: the sentence should be revised to “Enhance the symbol error rate” rather than “decrease the SER”, which seems contradictory.
• In Figure 5c, the distinction between the top and bottom plots is unclear. Please clarify.
• The legend of the same figure contains an incorrect reference: it should refer to Figure 4, not Figure 3.
• There are also some unnecessary repetitions throughout the text.

Beyond these formal remarks, while the section on device control using brain signals is reasonably well explained, the part related to encrypted communication remains highly unclear. The role of the human in this encryption process is not clearly defined. As it stands, a pseudorandom VSK1 would likely perform better without human involvement.

Moreover even in this resubmission many terms remain vague, such as “fusing” and “interaction” in Figure 2, and even

within the core of the article.

Actually, the links between the components sometimes appear artificial — for example, the LED blinking, supposedly tied to BSTCM, could just as easily be decoupled and displayed on a conventional screen not connected to the reconfigurable surface.

In summary, while the revised version is more readable, the work still lacks of clarity and overall of novelty besides the integration of diverse technologies: brain signal acquisition, ML, the use of RIS, and most notably, secure wireless communication. This interdisciplinary approach is commendable. However I am not sure it is sufficent for an article in nature comms

Authors' Responses

Firstly, we appreciate all reviewers for the constructive and insightful comments and suggestions, which helped us improve the quality of our manuscript significantly. Based on these comments and suggestions, we revised the manuscript very carefully. Below are our point-by-point responses to the reviewers' comments.

REVIEWER #1

We would like to thank **Reviewer #1 (R1)** for the positive comment regarding to our manuscript. In particular, **R1** states that "*The proposed work involves a brain-machine interface combined with space-time coding metasurfaces for wireless communication. However, the core novelty of the research is unclear, as it primarily seems to focus on combining EEG data with metasurfaces. Additionally, the paper does not address security measures for the wireless communication aspect, and providing quantitative metrics, such as bit error rate (BER) for the wireless links, would be beneficial. Moreover, there are several typos in the article, so it is recommended that the paper be thoroughly proofread.*" In addition, **R1** provides several suggestions on how to improve our manuscript. We thank **R1** for the positive feedback on our manuscript and do appreciate the suggestion. The response to **R1**'s suggestion and the change we have made in response to this comment is outlined below.

Comment 1.1: *The proposed work involves a brain-machine interface combined with space-time coding metasurfaces for wireless communication. However, the core novelty of the research is unclear, as it primarily seems to focus on combining EEG data with metasurfaces.*

Response: We sincerely appreciate your insightful comments. We acknowledge that our original manuscript did not adequately elaborate on the core novelty. The core novelty of this paper is the innovative **fusion-coding method between SSVEP visual stimulation coding and metasurface space-time coding (STC) for secure information interaction at the physical layer**. To clearly demonstrate the core novelty, we have revised Figure 1 (as shown in Figure R1 below) and listed the

specific illustration point-by-point.

- **Fusion coding scheme.**

In most BCI systems, users are typically presented with visual stimuli generated by computer screens, then corresponding EEG signals are acquired and recognized to interpret user intent [R1-R4]. Subsequently, the corresponding control commands are transmitted to operate peripheral devices. However, traditional approaches face significant limitations in enabling the simultaneous integration of visual stimulation and real-time information interaction, thereby restricting efficiency and making it difficult to deal with complex task demands. We address this challenge with a deep fusion coding scheme that combines SSVEP visual stimulation coding with metasurface STC to control the BSTCM for simultaneous visual stimulation and information interaction. This fusion enables the system to not only enhance efficiency but also perform advanced EM wave manipulation across both spatial and frequency domains, promoting a wide range of applications.

- **An integrated platform for both SSVEP and EM wave modulation.**

Various visual stimulators have been employed to evoke SSVEP components, including computer screens and LEDs. Each of these stimulators has been investigated in implementing the SSVEP-based BCI system [R5-R10]. However, the visual stimuli for BCI provided by display screens or LEDs are usually separate from the interaction and control devices, which increases system complexity. As depicted in Figure R1, **the presented BSTCM system integrates visual stimulation and EM wave manipulation by embedding LEDs onto an STC metasurface**, thereby achieving a combination of visual stimulators and control devices into a more integrated platform. This novel integrated platform renders the entire system more compact and lightweight, enhancing both its mobility and portability. The proposed integrated platform reduces the risk of connection failures by reducing external connections and physical interfaces, thereby improving the overall reliability of the system. The integration of SSVEP's frequency characteristics with the harmonic capabilities of STC metasurface allows for flexible EM wave manipulation, surpassing the previous programmable metasurfaces combining with BCI interface. [R11, R12].

- **A physical-layer secure platform.**

Physical layer security leverages the uniqueness and reciprocity of physical channels to achieve information encryption, which relys on the physical functionality of the BSTCM system. Multiple

harmonic-frequency-dependent physical channels can be constructed to realize multichannel information transmission because the BSTCM can control the propagation direction and harmonic power distribution in both the spatial and frequency domains [R13-R16]. Consequently, we further propose a secure encrypted wireless communication system that ensures reliable and secure information exchange between the human brain and external devices. Additionally, this innovative platform enables the remote control of smart devices based on human intention, offering a versatile and scalable solution for future brain-computer interaction systems.

References:

[R1] M. C, X. R. Gao, S. K. Gao and D.F. Xu, Design and implementation of a brain-computer interface with high transfer rates, *IEEE Transactions on Biomedical Engineering,* vol. 49, no. 10, pp. 1181-1186, Oct. 2002.

[R2] X. R. Gao, D. F. Xu, M. C, and S. K. Gao, A BCI-based environmental controller for the motion-disabled, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 137-140, June 2003.

[R3] B. Rebsamen et al., A Brain Controlled Wheelchair to Navigate in Familiar Environments, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 6, pp. 590-598, Dec. 2010.

[R4] Y. He, D. Eguren, J. M. Azorín, R. G. Grossman, T. P. Luu, and J. L. Contreras, Contreras-Vidal, J. L. Brain-Machine Interfaces for Controlling Lower-Limb Powered Robotic Systems. *J. Neural Eng.* 2018, 15 (2),
021004.

[R5] Z.H. Wu, Y.X. Lai, Y. Xia, D.Z. Yao, Stimulator selection in SSVEP-based BCI, *Medical Engineering & Physics* 30 (2008) 1079–1088.

[R6] D.H Zhu, Jordi Bieger, Garcia Molina, Gary, Aarts, Ronald M., A Survey of Stimulation Methods Used in SSVEP-Based BCIs, *Computational Intelligence and Neuroscience,* 2010, 702357, 12 pages, 2010.

[R7] R. Na, et al. An embedded lightweight SSVEP-BCI electric wheelchair with hybrid stimulator. *Digit. Signal Process*. 116, 103101 (2021).

[R8] O. B. Guney, M. Oblokulov, and H. Ozkan, A Deep Neural Network for SSVEP-Based Brain-

Computer Interfaces, *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 2, pp. 932-944, Feb. 2022.

[R9] J. Mu, D. B. Grayden, Y. Tan and D. Oetomo, Comparison of Steady-State Visual Evoked Potential (SSVEP) with LCD vs. LED Stimulation, *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, 2020, pp. 2946-2949.

[R10] X. Chen, X. Huang, Y. Wang and X. Gao, Combination of Augmented Reality Based Brain-Computer Interface and Computer Vision for High-Level Control of a Robotic Arm, I*EEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 12, pp. 3140-3147, Dec. 2020.

[R11] Ma, Q., Gao, W., Xiao, Q. et al. Dire ctly wireless communication of human minds via non-invasive brain-computer-metasurface platform. *eLight* 2, 11 (2022).

[R12] Q. Xiao, et al. Electromagnetic Brain–Computer–Metasurface Holography. *ACS Photonics* 10, 2249–2256 (2023).

[R13] L. Zhang, et al. Space-time-coding digital metasurfaces. *Nat. Commun.* 9, 4334 (2018).

[R14] L. Zhang, et al. A wireless communication scheme based on space- and frequency-division multiplexing using digital metasurfaces. *Nat. Electron.* 4, 218–227 (2021).

[R15] G.-B. Wu, J. Y. Dai, Q. Cheng, T. J. Cui, and C. H. Chan, Sideband-free space–time-coding metasurface antennas. *Nat. Electron*. 5, 808–819 (2022).

[R16] L. Zhang, and T. J. Cui, Space-Time-Coding Digital Metasurfaces: Principles and Applications. *Research* 2021, 2021/9802673 (2021).

Action Taken: Following your comment, we have supplemented a figure to demonstrate the core novelty in Figure 1. To emphasize this core novelty, we have made careful revisions to the abstract and the main text. This adjustment enhances the clarity and readability of our manuscript.

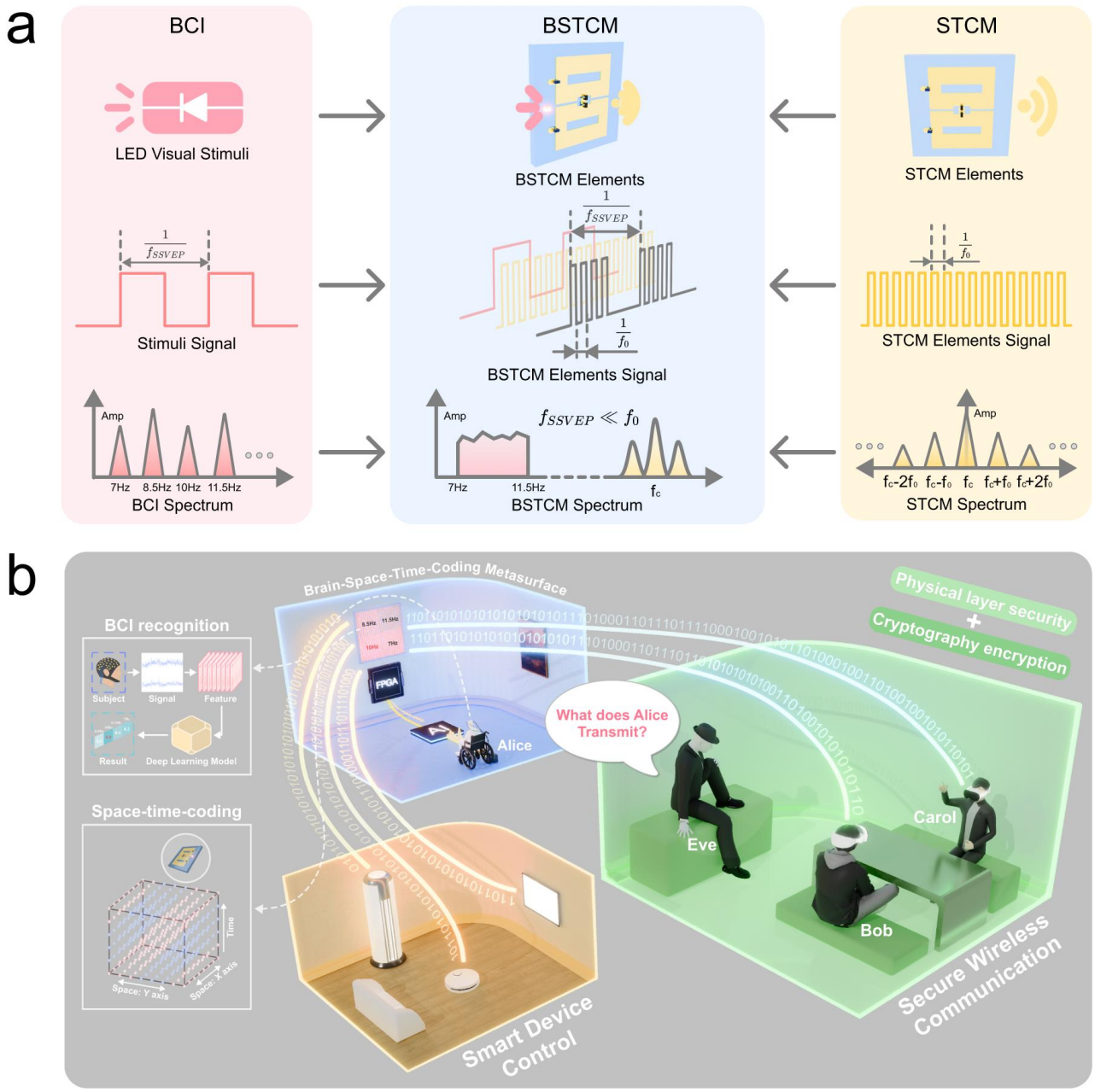

Figure R1. The schematical diagram of the presented BSTCM. a. The schematic diagram of fusing visual stimulation coding with metasurface coding. The STC metasurface can correctly support the visual stimuli for the SSVEP-based BCI and facilitate the information interaction with the external environment. **b.** The BSTCM system integrates the STC metasurface into the SSVEP-based BCI systems. Based on the BSTCM system, a secure encrypted wireless communication system is realized by combining with the variant HSKs and VSS method. In this way, smart devices can be controlled by the human mind.

[Abstract]

…However, most existing BCI technologies are based on simple signal transmission and independent of other interface devices, with limited consideration for the reliability and security of human brain's information interaction in complicated wireless environments. To address the formidable limitation, we propose a deep fusion coding scheme that combines the BCI visual stimulation coding with metasurface space-time coding at the physical layer, enabling reliable and secure information transfers between the human brain and external devices. To achieve this fusion coding scheme, we present a brain space-time-coding metasurface platform to implement a secure wireless communication system by using harmonic-encrypted beams. We design and fabricate a proof-of-principle prototype and experimentally show that the proposed wireless BCI scheme can establish a remote but safeguarded paradigm for human-machine interactions and intelligent metasurfaces, providing a potential direction in the future secure wireless communications.

[Introduction]

Here, we propose a deep fusion coding scheme that combines SSVEP visual stimulation coding with metasurface coding for secure information interactions at the physical layer. To implement this scheme, we present a brain space-time-coding metasurface (BSTCM) platform that seamlessly integrates the visual stimulation for SSVEP-based BCIs with the information interaction to external environment, significantly improving the system compactness and reliability. To enhance BCI security, we propose a secure encrypted wireless communication system that integrates the physical layer security with the cryptography encryption method. Specifically, the encryption method combines variant harmonic-based keys (HSKs) with visual secret sharing (VSS) technology. Experimental results demonstrate high bit error rate (BER) of nearly 50% for the eavesdroppers and the secrecy capacity of approximately 1.9 dB, validating the security of the encrypted wireless communication system. Finally, smart device control is presented by using the BSTCM system, realizing intelligent human-machine interactions. The proposed BSTCM can establish a new paradigm of the human-machine interactions and secure wireless communications.

[System configuration]

The proposed deep fusion coding method combines the SSVEP visual stimulation coding with the metasurface coding for secure information interactions at the physical layer, as presented in Fig. 1a. The SSVEP components in user's brain signals can be induced by low-frequency flickering visual stimuli (4~50Hz), while the STC metasurface produces a series of harmonic frequencies, revealing notable similarities in the frequency domain. Given that the switching frequency of the metasurface STC is significantly higher than the frequency of flickering stimulus, it becomes feasible to embed the STC operations in the high-level intervals of the low-frequency flickering stimuli. This integration enables seamless fusion of the two coding mechanisms, as illustrated in Fig. 1a. To realize this fusion coding method, an integrated BSTCM platform for both SSVEP and EM-wave modulation is schematically demonstrated in Fig. 1, which consists of an STC metasurface, an SSVEP-based BCI, and an FPGA control module. The metasurface element contains a meta-structure to regulate the EM wave, and a light-emitting diode (LED) stimulator for the SSVEP paradigm.

The STC metasurface is divided into four LED flickering partitions, operating at different

frequencies: 8.5Hz, 10Hz, 11.5Hz, and 7Hz. The corresponding SSVEP-based EEG signals can be accurately extracted by a head-worn EEG cap when the operator is exposed to the LED flickering stimuli at distinct frequencies with the assistance of FPGA. To improve the efficiency and accuracy of the EEG signal identification, we propose a deep learning model to recognize the extracted EEG signals, as shown in Fig. 3c. After completing the identification process, the recognized EEG signals are classified as different commands and then transmitted to FPGA, which fuses the metasurface coding into SSVEP visual stimulation coding to generate a specific BSTCM control signal. This allows for the simultaneous accomplishment of both visual stimulation and EM-wave modulation by the BSTCM signal, thereby enabling a wide range of potential applications. The BSTCM system empowers individuals to customize and modulate the EM waves by the human mind, enabling a variety of EM functions and enhancing the flexibility and interactivity of human-machine systems.

[Conclusions]

We presented a deep fusion coding scheme that combines the visual stimulation coding with the metasurface coding, and developed an integrated BSTCM platform to enable simultaneous visual stimulation and EM manipulation. The proposed system harnesses the human brain intelligence and the flexible EM-control capabilities of the STC metasurface, enabling the information interaction within the EM field to establish reliable and stable communication environments. Furthermore, we demonstrated a secure encrypted wireless communication system that combines the physical layer security and the cryptography encryption method, greatly enhancing the system security. Finally, remote control of smart devices was illustrated by the BSTCM system, realizing intelligent human-machine interactions. The proposed BSTCM, integrating the EM manipulation capability with the intelligence of the human brain, provides a new paradigm for the interactions among human-machine interfaces and the future 6G wireless communication applications.

Comment 1.2: *Additionally, the paper does not address security measures for the wireless communication aspect, and providing quantitative metrics, such as bit error rate (BER) for the wireless links, would be beneficial.*

Response: Thank you for your professional comments. We have measured the corresponding security performance of the system [R17]. As shown in Figures R2a and b, the system security metrics are measured when the illegal user eavesdrops separately from the spatial and frequency domains, respectively. For spatial-domain eavesdropping, the eavesdropper is positioned at different angles relative to the intended transmission direction, with angles ranging from 35° to 60°. For frequency-domain eavesdropping, eavesdropping attempts are conducted at different harmonic frequencies, including -2^{nd}, -1^{st}, 0^{th}, $+1^{st}$, and $+2^{nd}$ orders, while the eavesdropper is fixed at a spatial location of

35°. All measured results are demonstrated in Figures R2c and d. The measured bit error rate (BER) results are nearly 50% when the illegal user with -1st harmonic frequency is eavesdropping in different spatial locations from 35° to 60°. Meanwhile, the illegal users with different harmonic frequencies from -2nd to +2nd order have been tested when they are located in the 35° spatial direction. The corresponding BERs are nearly 50%. Through the excellent manipulation of harmonic beams by the BSTCM system, the signal-to-noise ratio (SNR) for legal users is measured as approximately 22 dB, ensuring reliable and high-quality communication. In contrast, the SNR for illegal users located in the 35° spatial direction is less than 5 dB, effectively degrading the quality of intercepted signals. Based on these SNR measurements, the secrecy capacity ($C_s = \log_2(1+S_{legal}/N_{legal}) - \log_2(1+S_{illegal}/N_{illegal})$) between the legal user and the illegal user is calculated as about 1.9 dB, which strongly supports the system's capability to secure information at the physical layer. Furthermore, the integration of spatial and frequency-dependent security mechanisms within the BSTCM system establishes robust physical-layer security. The proposed dual-layer encrypted wireless communication approach not only ensures the confidentiality of transmitted data but also significantly increases the difficulty of unauthorized access, even under sophisticated attack scenarios. The experimental results further substantiate the effectiveness of the BSTCM-based wireless communication system in achieving secure and reliable information transmission in complex communication environments.

References:

[R17] Venkatesh, S., Lu, X., Tang, B. et al. Secure space–time-modulated millimetre-wave wireless links that are resilient to distributed eavesdropper attacks. *Nat Electron* 4, 827–836 (2021).

Action Taken: Following your comment, we have added the security performance discussions in the main text and supplementary materials of the revised manuscript. We sincerely appreciate **R1** for raising such a valuable question, as it has greatly contributed to enhancing the quality of our work.

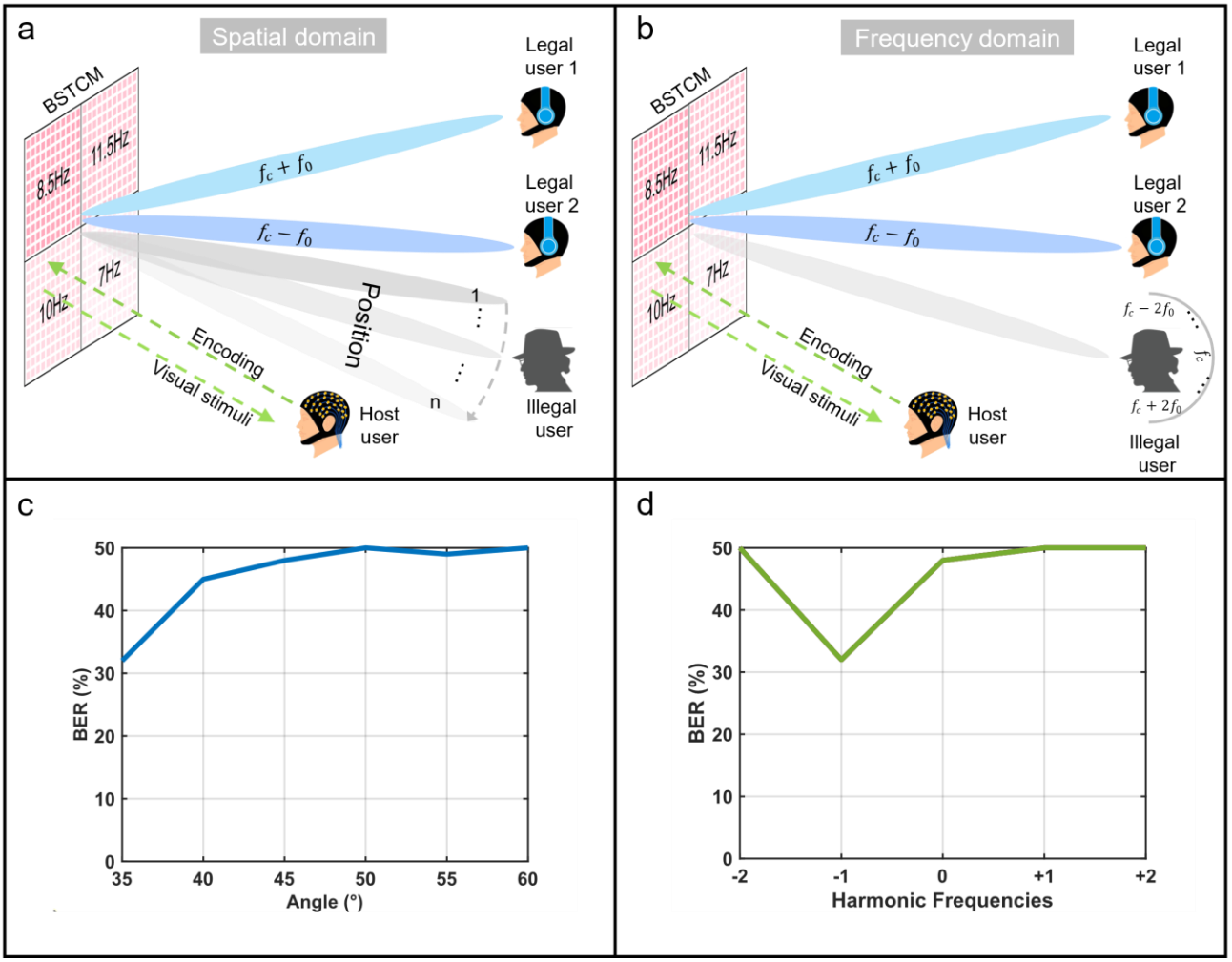

Figure R2. The security measurement of the proposed secure wireless communication scheme. a. The diagram of security test scenarios for illegal users in different spatial locations from 35° to 60°. b. The diagram of security test scenarios for illegal users in different harmonic frequencies from -2nd to +2nd order. c. The measured results for illegal users in different spatial locations from 35° to 60°. d. The measured results for illegal users in different harmonic frequencies from -2nd to +2nd order.

[Introduction]

> …Experimental results demonstrate high bit error rate (BER) of nearly 50% for the eavesdroppers and the secrecy capacity of approximately 1.9 dB, validating the security of the encrypted wireless communication system….

[Secure encrypted wireless communication system]

> …To verify the security performance of the BSTCM system, we measured the signal-to-noise ratio (SNR) and BER results. The SNR for two legal users is measured as approximately 22 dB, while the SNR for illegal user is less than 5 dB. The security capacity is calculated as approximately 1.9 dB. Furthermore, the BERs are measured as nearly 50% when the illegal user eavesdrops separately from the spatial and frequency domains (see Supplementary Note 15 for additional details). These

results validate the robustness and effectiveness of the proposed system in ensuring high level of security and confidentiality, demonstrating its resilience against unauthorized interception and its capability to safeguard the secret information during transmission.

[Supplementary Note 15: The Security Performance of the Proposed Secure Wireless Communication]

We have presented a cryptography encryption method as described above, which combines the visual secret sharing (VSS) method and the variant harmonic-based secret keys. The proposed scheme distributes confidential information between two users, ensuring that interception of a single communication channel does not disclose any information of the shared secret. In BSTCM-assisted wireless communication systems, multiple harmonic-frequency-dependent physical channels can be constructed to realize multichannel information transmission because the BSTCM can control the propagation direction and harmonic power distribution of the EM waves in both the spatial and frequency domains, which offers the physical layer security that leverages the characteristics of the physical channel to achieve information encryption. To verify the security of the proposed scheme, we have tested the signal-to-noise ratio (SNR) and BER of the BSTCM system. As shown in Supplementary Figure 26a and b, the system security is measured when the illegal user eavesdrops separately from the spatial and frequency domains, respectively. Firstly, we measured the BER results of the illegal user with -1st harmonic frequency in different spatial locations from 35° to 60° in the BSTCM-based wireless communication system. The BERs are nearly 50% shown in Supplementary Figure 26c. The illegal users with different harmonic frequencies from -2nd to +2nd order have been tested when they are located at 35° spatial direction. Supplementary Figure 26c demonstrated that BERs are nearly 50%. In addition, the SNR for the two legal users is measured as approximately 22 dB while the illegal user is less than 5 dB. The secrecy capacity is denoted as the difference in the channel capacity between the legal user and the illegal user. It is described as follows: $C_s = log_2(1 + S_{illegal}/N_{legal}) - log_2(1 + S_{illegal}/N_{illegal})$, where S_{legal} and $S_{illegal}$ are the powers received by the legal user and the illegal user and N_{legal} and $N_{illegal}$ denote the corresponding noise. The secrecy capacity of our system is calculated as about 1.9 dB. Consequently, the proposed scheme is highly secure and not easily cracked unless he/she can crack the channel direction, harmonic frequency, and two HSKs simultaneously.

Comment 1.3: *Moreover, there are several typos in the article, so it is recommended that the paper be thoroughly proofread.*

Response and Action Taken: Thank you for raising the question. We have thoroughly proofread the manuscript and corrected all identified typographical errors. We greatly appreciate your suggestion, as it has helped improve the quality of the manuscript.

[Abstract]

…However, most existing BCI technologies are based on simple signal transmission and independent of other interface devices, with limited consideration for the reliability and security of human brain's information interaction in complicated wireless environments….

[Introduction]

…The frequency-dependent SSVEP responses and programmable harmonic characteristics of space-time-coding (STC) metasurfaces have notable similarities….
…Hence, the STC metasurface is a potential candidate for deep information modulation and interaction in the SSVEP-based BCI system owing to its capability to flexibly manipulate the EM waves and interact with the frequency-dependent visual stimulation.

[System configuration]

…After completing the identification process, the recognized EEG signals are classified as different commands and then transmitted to the FPGA, which fuses the metasurface coding into SSVEP visual stimulation coding to generate a specific BSTCM control signal….
…In the encrypted wireless communication system, the harmonic frequencies are adopted as the secret keys and the VSS method is employed to encrypt the target information[46-48]….

[SSVEP signal recognition]

…Subsequent application of fast Fourier transform (FFT) to these sub-bands reveals their amplitude spectra, highlighting the frequency components of the SSVEP signals….

[Design of STC metasurface]

…The transmission of digital symbols "0" and "1" is achieved by encoding a sequence of repeated data frames, with every set of 4 frames serving as the synchronous frames….

[Mind-controlled smart devices]

…As displayed in Figs. 6c and d, the measured radiation patterns exhibit good consistency with theoretical predictions corresponding to the STC matrices….

REVIEWER #2

We thank **Reviewer #2** (**R2**) for the positive feedback regarding our manuscript. In particular, **R2** states that "*The experiments presented are compelling and impressive. Overall, this work is fancy and merits acceptance. There are some optional suggestions to enhance the manuscript: The introduction section contains a considerable amount of detail, particularly in the last paragraph. For clarity, the authors are kindly suggested that the authors condense this section.*" In addition, **R2** has given several insightful and valuable comments, which are very helpful to further improve the quality of our manuscript. We have addressed all comments carefully, and all changes are listed as follows.

Comment 2.1: *This work integrates brain-computer interfaces with spatiotemporal metasurfaces, which is quite innovative. A harmonic encryption method is employed to achieve secure wireless communications. Although the combination of brain-computer interfaces and metasurfaces has been explored for several years, most of the researches focused on static metasurfaces. The application of spatiotemporal metasurfaces introduces various harmonics, which can be utilized for encryption and wireless communications, as demonstrated in the main text. The experiments presented are compelling and impressive. Overall, this work is fancy and merits acceptance. There are some optional suggestions to enhance the manuscript: The introduction section contains a considerable amount of detail, particularly in the last paragraph. For clarity, the authors are kindly suggested that the authors condense this section.*

Response and Action Taken: Thank you for your professional suggestions. We have modified the Introduction part. We thank **R2** for the professional comment, which is very helpful in enhancing the quality of our manuscript.

[Introduction]

> …Wireless transmission of brain signals in the BCI systems is vulnerable to theft and attacks, potentially leading to inaccurate control commands and unauthorized privacy breaches. Although some methodologies have been proposed to enhance the security and privacy in BCIs[23-25], the encryption mechanisms specifically tailored to these systems remain largely unexplored. Moreover, visual stimulation is often isolated from back-end information processing, lacking deep integration and interaction. With the increasing demand for secure BCI systems, it is essential to develop intelligent interactions in secure and reliable communication environment.… Therefore, the STC metasurfaces can be used as a promising method that not only provides visual stimulation but also enhances the security of the BCI systems at the physical layer owing to their powerful ability to flexibly manipulate electromagnetic (EM) waves in both time and space domains.

[Introduction]

Here, we propose a deep fusion coding scheme that combines SSVEP visual stimulation coding with metasurface coding for secure information interactions at the physical layer. To implement this scheme, we present a brain space-time-coding metasurface (BSTCM) platform that seamlessly integrates the visual stimulation for SSVEP-based BCIs with the information interaction to external environment, significantly improving the system compactness and reliability. To enhance BCI security, we propose a secure encrypted wireless communication system that integrates the physical layer security with the cryptography encryption method. Specifically, the encryption method combines variant harmonic-based keys (HSKs) with visual secret sharing (VSS) technology. Experimental results demonstrate high bit error rate (BER) of nearly 50% for the eavesdroppers and the secrecy capacity of approximately 1.9 dB, validating the security of the encrypted wireless communication system. Finally, smart device control is presented by using the BSTCM system, realizing intelligent human-machine interactions. The proposed BSTCM can establish a new paradigm of the human-machine interactions and secure wireless communications.

Reviewer #3

We sincerely appreciate **Reviewer #3** (**R3**) for the time and effort invested in reviewing our manuscript and offering a wealth of insightful comments and thoughtful suggestions. These comments and suggestions are extremely valuable to enhance the quality of our manuscript. Following **R3**'s constructive guidance, we have carefully revised the paper to improve its clarity and comprehensibility. The authors thank **R3** for the invaluable comments and constructive suggestions, which have greatly contributed to significantly improving the quality of this paper.

Comment 3.1: *The article is very difficult to read. It took me an entire afternoon to go through it, paragraph by paragraph, while also referring to the supplemental materials. Despite this effort, many aspects remain unclear to me. I understand the overall purpose of the work and some individual steps, but I do not have a comprehensive understanding of the process involved in achieving secure wireless communications.*

Response: Thank you for your professional comments. The proposed high-security model of the BSTCM-based wireless communication system is shown in Figure R3, which integrates physical layer security and cryptography encryption method. To clearly demonstrate the proposed secure wireless communication process, we have answered the specific questions point-by-point and clarified the secure wireless communication process of our work.

Physical layer security. Physical layer security leverages the uniqueness and reciprocity of physical channels to achieve information encryption. In BSTCM-assisted wireless communication systems, the physical channel represents the physical functionality of the BSTCM, which can be modulated to transmit information by spatial beamforming. Multiple harmonic-frequency-dependent physical channels can be constructed to realize multichannel information transmission because the STC metasurface can control the propagation direction and harmonic power distribution of the EM waves in both the spatial and frequency domains [R1-R5]. This makes the encryption process less susceptible to attacks such as brute force decryption or algorithmic weaknesses because the system security is integrated into the frequency-dependent physical channels. The detailed steps are listed as below:

1) **Step 1:** The process begins with the optimization of two STC matrices using the BPSO algorithm,

each corresponding to distinct harmonic frequencies and spatial directions (see Supplementary Note 8). Specifically, one STC matrix is tailored for the $+1^{st}$ harmonic frequency at the specified direction of -15°, while the other is optimized for the -1^{st} harmonic frequency at the specified direction of 30°. Two optimized STC matrices are crucial for constructing secure and directional transmission physical channels. Once optimized, two STC matrices are pre-stored within the FPGA module of the BSTCM platform. The FPGA serves as a control and processing module, enabling seamless fusion between the visual stimulation coding and two STC matrices. This configuration effectively establishes two distinct physical channels, each characterized by its unique harmonic frequency and directional specificity. These two physical channels allow for reliable information delivery to designated users, as the $+1^{st}$ harmonic channel (-15°) and -1^{st} harmonic channel (30°) are spatially and spectrally isolated.

2) **Step 2:** The transmission of two encrypted information begins with the operator gazing at one of the designated regions on the BSTCM platform. Each region corresponds to a specific LED flickering frequency, which induces distinct SSVEPs in the operator's brain signals. These brain signals, unique to the gazed frequency, are captured by the EEG device and processed through a deep learning model for accurate decoding and classification into specific commands. Then, the commands are transformed into a BSTCM control signal, which simultaneously encodes the operator's encrypted information while applying advanced EM wave manipulations. Two encrypted messages are transmitted through two physical channels with specific spatial directions and harmonic frequencies. This approach ensures that the transmission process is resilient against potential eavesdropping or interference, providing both robustness and security.
3) **Step 3:** The encrypted information is carefully transmitted to two legal users positioned at specified spatial directions of -15° and 30°, corresponding to the $+1^{st}$ and -1^{st} harmonic frequencies, as shown in Figure R3b. As the encrypted information propagates at two physical channels, the combination of spatially directed beams and frequency-specific harmonics allows the intended recipients to accurately decode the transmitted information. The legal users can extract the encoded data by tuning their receivers to the appropriate harmonic frequency and spatial direction, ensuring a secure and reliable communication link. Meanwhile, the integrated spatial and frequency-domain physical channels impose significant barriers for illegal users attempting to intercept the transmission. The precise alignment of the BSTCM signal with specific spatial angles

and harmonic frequencies means that any deviation from these parameters results in signal degradation or complete failure of interception. This dual-channel security mechanism effectively safeguards the transmitted information, making the BSTCM system highly resilient against potential eavesdropping and unauthorized access.

Cryptography encryption method. Based on our proposed BSTCM platform, we present a new cryptography encryption method, which integrates the visual secret sharing (VSS) method and the variant harmonic-based secret keys. The detailed security process is demonstrated in Figure R3a and Figure R7.

1) **Step 1:** The transmitting information (*M*) can be encrypted into two ciphertexts (*W1* and *W2*) using two same-length harmonic secret keys (*K1* and *K2*) according to the encryption process (see Supplementary Note 12). As shown in Figure R4, the secret image "H" is a grayscale matrix composed of 5×5 black pixels (digital "1") and white pixels (digital "0"). The encryption coding scheme is presented to encrypt the secret information. Two shared sub-pixels of the secret pixel representing digital "1" are randomly encoded into $+1^{st}$ and -1^{st} harmonic frequencies. Conversely, for a secret pixel of digital "0", two same sub-pixels are encoded into either $+1^{st}$ or -1^{st} harmonic frequencies. Hence, each pixel of the secret image "H" is encrypted into two shared sub-pixels, which are ultimately combined to form two VSKs (*W1* and *W2*). Importantly, neither VSK1 nor VSK2 alone contains sufficient information to reveal the original secret information. In order to verify the effectiveness of the cryptography encryption method, Figure R5 also shows the encryption process of other secret images.
2) **Step 2:** Harmonic frequencies corresponding to two shared sub-pixels are combined into two harmonic HSKs (*K1* and *K2*), which are securely shared with two designated legal users. These HSKs form the basis of the encoding process, ensuring that only legal users can decrypt the transmitted information. Using the BSTCM platform, two VSKs (*W1* and *W2*) are encoded into information (*X1* and *X2*). The encoding process integrates the VSKs with the harmonic frequencies of the STC metasurface, which modulates the information to propagate through two distinct physical communication channels. These channels are differentiated by their respective harmonic frequencies and spatial directions, corresponding to the specified angles of -15° and 30°, as shown in Figure R3b. As the encoded information (*X1* and *X2*) propagate through their respective

physical channels, the characteristics of the communication environment, such as channel noise, interference, and propagation losses, affect the transmitted signals. These effects cause the original information *X* to transform into the received information *Y* by the legal users.

3) **Step 3:** After receiving and processing the transmitted signals, each legal user decodes their respective received signal (*Y1* and *Y2*) into the encrypted information (*W1* and *W2*). The schematical decryption process is shown in Figure R6. The two legal users, utilizing their shared HSKs (*K1* and *K2*), combine their decoded components (*W1* and *W2*) through the proposed decryption scheme to reconstruct the original transmitted message *M*. Illegal users, lacking access to the correct HSKs and the required spatial and frequency alignment, are unable to decrypt the received signals or reconstruct the target information. The BSTCM system's robust encoding mechanisms enable the target users to accurately reconstruct the original information using the shared HSKs (*K1* and *K2*) and the spatial and frequency-domain alignment provided by the STC metasurface. This collaborative decryption process ensures that both users contribute to the final information reconstruction, emphasizing the cooperative nature of the secure communication protocol.

We have presented a secure wireless communication system by the BSTCM platform by combining the physical layer security and cryptography encryption method. The illegal user is unable to acquire any meaningful information without access to the direction and harmonic-frequency information from a single physical communication channel, thus improving the security of wireless communication.

References:

[R1] M. Cui, G. Zhang and R. Zhang, Secure Wireless Communication via Intelligent Reflecting Surface. *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1410-1414, Oct. 2019.

[R2] Poor HV, Schaefer RF. Wireless physical layer security. *Proc Natl Acad Sci USA*. 2017 Jan 3;114(1):19-26.

[R3] Y. Zheng, K. Chen, Z. Xu, N. Zhang, J. Wang, J. Zhao, Y. Feng, Metasurface-Assisted Wireless Communication with Physical Level Information Encryption. *Adv. Sci*. 2022, 9, 2204558.

[R4] Wei, M., Zhao, H., Galdi, V. et al. Metasurface-enabled smart wireless attacks at the physical layer. *Nat Electron* 6, 610–618 (2023).

[R5] Zhang, L., Chen, M.Z., Tang, W. et al. A wireless communication scheme based on space- and

frequency-division multiplexing using digital metasurfaces. *Nat Electron* 4, 218–227 (2021).

Action Taken: Following your comment, we have made some revisions in our manuscript. To clearly demonstrate the proposed secure wireless communication process, we have revised the secure wireless communication process and redrew the diagram again based on the content in Supplementary Note 12. We sincerely appreciate **R3** for raising such a valuable question, as it has greatly contributed to enhancing the quality of our work.

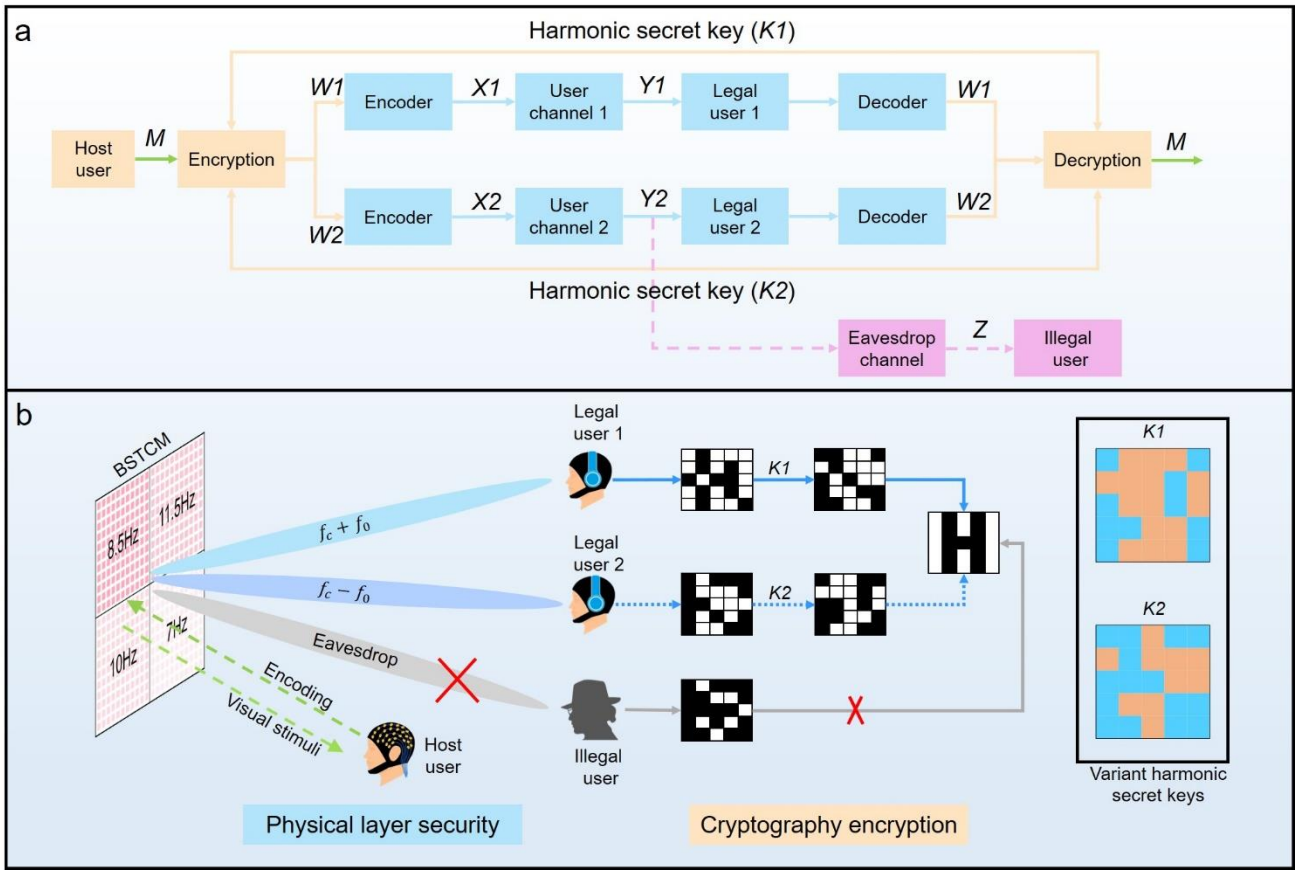

Figure R3. The proposed secure wireless communication scheme. a. The detailed flowchart of the proposed security model. b. The schematical diagram of the presented BSTCM-based secure wireless communication system.

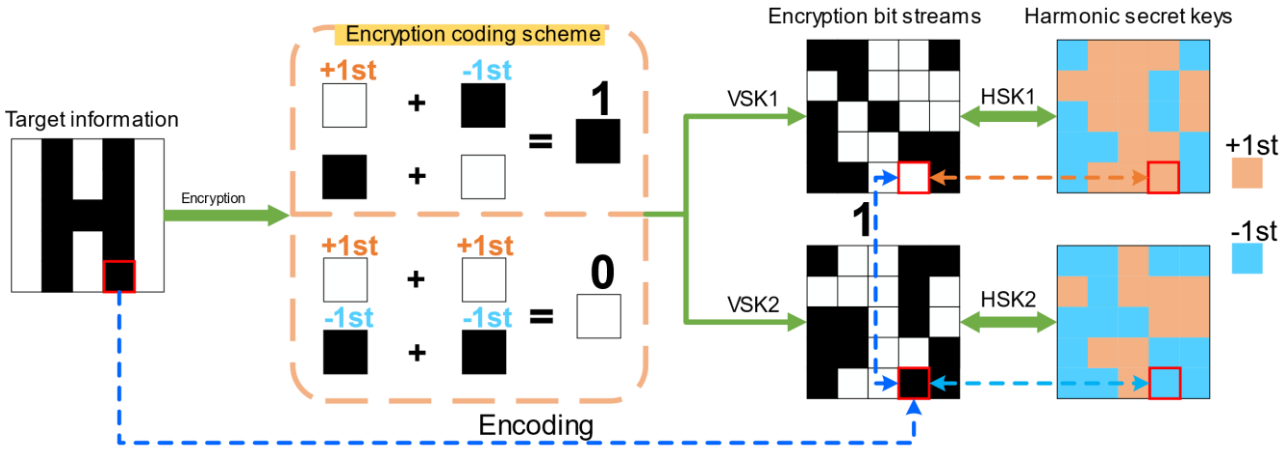

Figure R4. The encrypted encoding scheme for the secret image "H" combining the harmonic-secret keys and the VSS method.

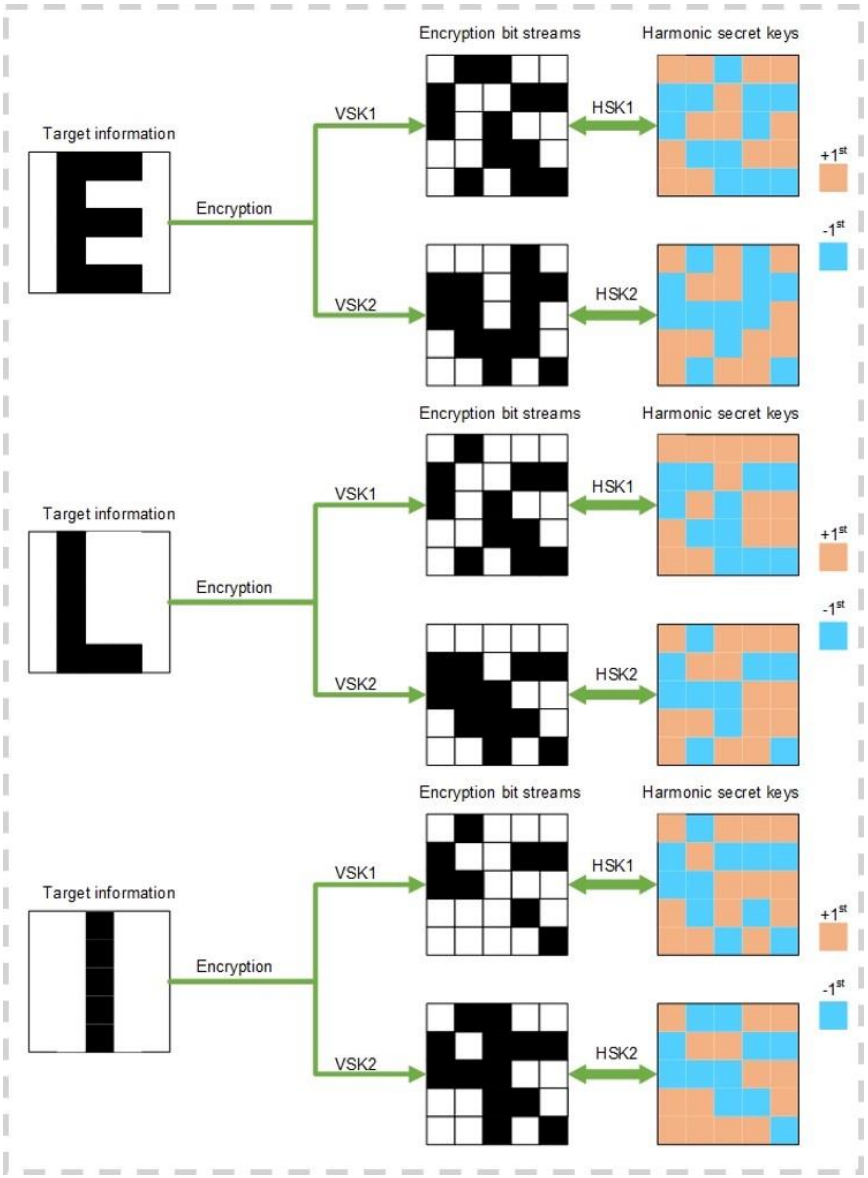

Figure R5. The encryption coding process of three secret letter images "E", "L", and "I". The accurate information is hidden in two VSKs associated with two HSKs.

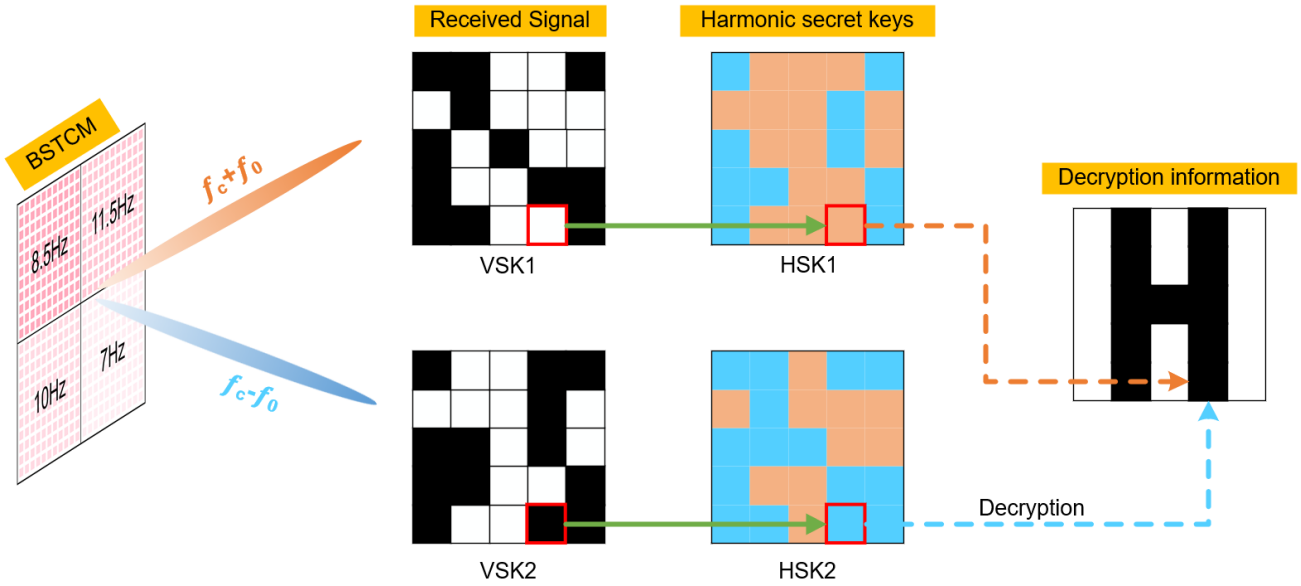

Figure R6. The decrypted coding process of the secret image "H" from two legal users at $+1^{st}$ and -1^{st} harmonic frequency according to two shared HSKs.

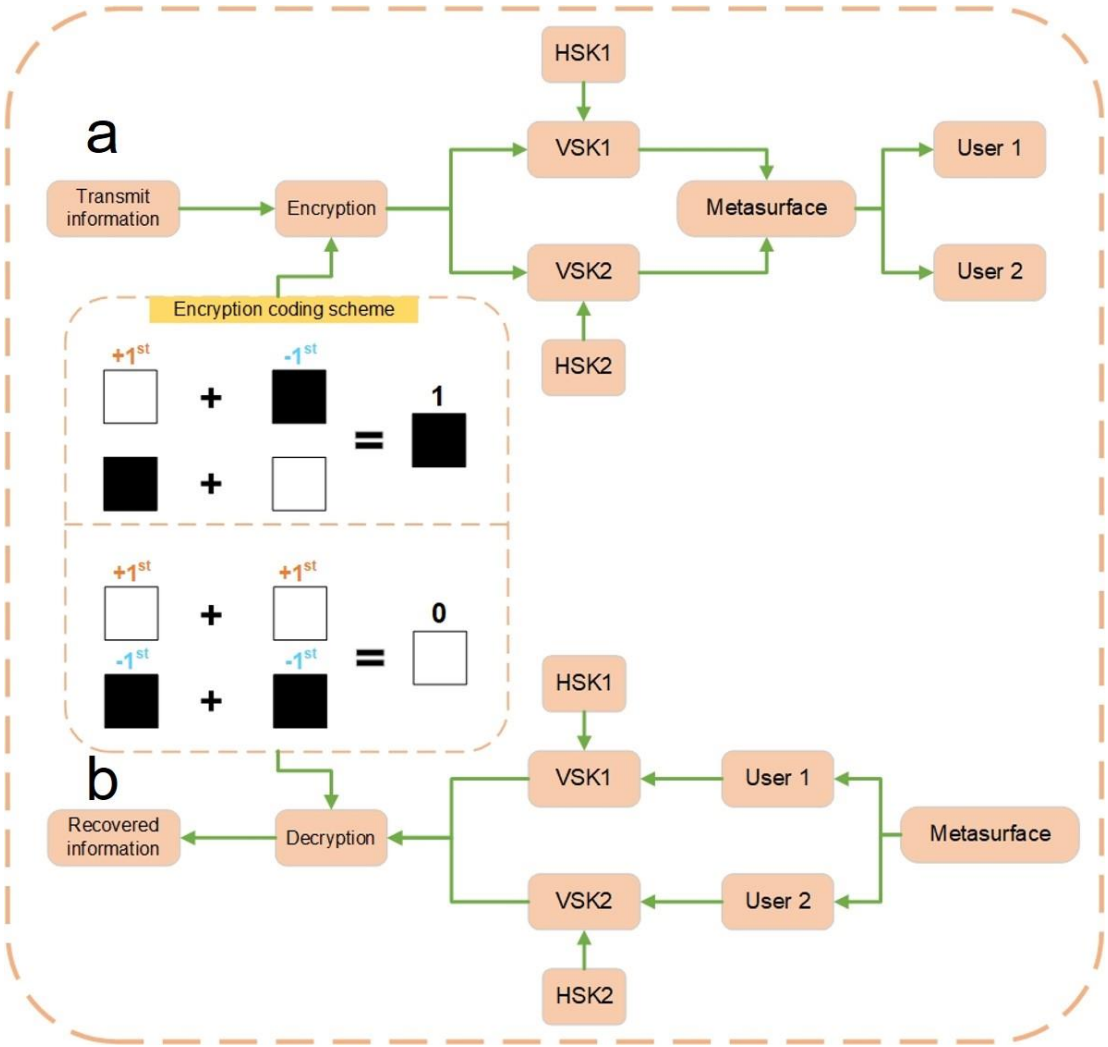

Figure R7. The presented wireless communication process. a. The target information is encrypted to send to two legal users. b. The target information is decrypted by the received information from two legal users.

[Design of STC metasurface]

…Hence, the good experimental results demonstrate that two frequency-dependent communication channels are constructed, providing a solid foundation for physical layer security.

[Secure encrypted wireless communication system]

…In the presented encryption strategy, the recovery of the original information is not achieved by merely stacking the two VSKs but relies on the corresponding HSKs. The encryption encoding scheme is shown in Fig. 5a. Two shared sub-pixels of the secret pixel representing digital "1" are randomly encoded into the +1st and -1st harmonic frequencies. Conversely, for a secret pixel of digital "0", two same sub-pixels are encoded into either the $+1^{st}$ or -1^{st} harmonic frequencies. As an illustrative example in Fig. 4a, the secret image "H", consisting of a grid of 5×5 black pixels (digital "1") and white pixels (digital "0"), is encrypted into two random VSK1 and VSK2 based on the presented encryption method. The harmonic frequencies associated with the two shared sub-pixels are further integrated into the corresponding HSKs, which are securely distributed to two legal users. Notably, the length of each HSK is equivalent to that of the corresponding VSKs, and HSKs dynamically adapt based on the content of VSKs. This dynamic variation enhances the system's security by ensuring that the HSK remains intrinsically tied to the encrypted VSK. Consequently, neither VSK1 nor VSK2 alone contains sufficient information to reveal the original secret information. By leveraging the proposed BSTCM system to construct two specific physical channels, the BCI operator (Alice) can sequentially transmit two bitstream sequences of two VSKs to two legal users (Bob and Carol) by the human mind….

[Supplementary Note 12: The harmonic-based encrypted wireless communication scheme]

…According to two HSKs, two VSKs will sent to two different users with different harmonic frequencies, respectively. As a result, every single secret is encoded into two matrices composed of random harmonic secret key sequences so that the effective information can be hidden as two VSK1 and VSK2. In this way, the contents of transmitting VSKs are completely random and meaningless, which makes either VSK1 or VSK2 strongly robust against different eavesdroppers. Then, two random VSK1/VSK2 are sent to two users with two harmonic frequencies by the constructed secure encrypted wireless communication system based on the BSTCM platform.
…Firstly, two horn antennas placed in a specific direction are used to receive messages from the host user sent through the BSTCM system….

[Supplementary Note 15: The Security Performance of the Proposed Secure Wireless Communication]

The proposed high-security model of the BSTCM-based wireless communication system is shown in Supplementary Figure 25, which integrates cryptographic encryption method and physical layer security. The detailed secure wireless communication process is demonstrated in Supplementary Figure 25a. The host user firstly selects the harmonic frequency orders to dynamically generate the harmonic secret keys (*K1, K2*) by randomly selecting -1^{st} and $+1^{st}$ harmonic modes and sharing them with two legal users. According to the encrypted process in Supplementary Figure 20, the target message (*M*) is firstly encrypted to two ciphertexts (*W1, W2*) using two keys. Two ciphertexts

(*W1, W2*) are encoded as the spatial information (*X1, X2*) and then sent to two secure user communication channels constructed from different harmonic frequencies and different spatial directions by the proposed BSTCM platform, as shown in Supplementary Figure 25b. Superimposed on the effects of channel characteristics, information *X* propagating through the channel transforms into *Y* received by the target users. Two legal users at specific directions can receive the corresponding information (*Y1, Y2*) and recover the target message (*W1, W2*) through two secret keys (*K1, K2*). Finally, all information from both two legal users is accurately decoded into the correct message M according to the decryption process in Supplementary Figure 20. However, the illegal user is unable to acquire any meaningful information without access to the direction and harmonic-frequency information. What's more, it is also impossible to crack the correct information from a single physical communication channel, greatly improving the security of information.

Comment 3.2: *The authors attempt to combine several popular techniques, such as reconfigurable intelligent surfaces, machine learning (ML), and the Metaverse. While I am not suggesting that this approach is without merit, it is very confusing, especially for readers who are not experts in secure wireless communications.*

Response: Thank you for your professional comments. We present an innovative BSTCM system through the embedding of LEDs onto an STC metasurface. More importantly, the fusion coding method deeply fuses SSVEP visual stimulation coding with the metasurface STC scheme to control the BSTCM for secure information interaction at the physical layer. Based on the presented BSTCM system, a secure encrypted wireless communication, which combines the physical layer security and cryptography encryption method, is presented to realize reliable and secure information transfer between the human brain and external devices. In addition, smart device control is presented by the BSTCM system, realizing intelligent human-machine interactions. **The STC metasurface** can control the propagation direction and harmonic power distribution of the EM waves in both the spatial and frequency domains, providing the physical layer security of the system. In our study, **the machine learning method** is mainly utilized to achieve brain signal classification from the extracted EEG signals, hence accurately obtaining the user's intention. A great number of published studies have demonstrated that the application of machine learning algorithms for recognizing EEG signals significantly enhances accuracy. The application of **BCI technology in the metaverse** holds immense potential, offering users a more immersive and interactive experience while facilitating seamless

integration between the virtual and real worlds [R6-R8]. Hence, we envision that our BSTCM system also may contribute to this field in the future. To avoid confusion, we have deleted the corresponding description.

References:

[R6] Liu, Y., Liu, R., Ge, J. & Wang, Y. Advancements in brain-machine interfaces for application in the metaverse. *Front Neurosci-switz* **18**, 1383319 (2024).

[R7] Zhu, H. Y., Hieu, N. Q., Hoang, D. T., Nguyen, D. N. & Lin, C.-T. A Human-Centric Metaverse Enabled by Brain-Computer Interface: A Survey. *IEEE Commun Surv Tut* **26**, 2120–2145 (2024).

[R8] Wu, D. et al. Virtual-Reality Interpromotion Technology for Metaverse: A Survey. *IEEE Internet Things* **10**, 15788–15809 (2023).

Action Taken: In response to **R3**'s question, we have revised the corresponding description and deleted the metaverse application in our manuscript.

[Abstract]

> …We design and fabricate a proof-of-principle prototype and experimentally show that the proposed wireless BCI scheme can establish a remote but safeguarded paradigm for human-machine interactions and intelligent metasurfaces, providing a potential direction in the future secure wireless communications.

[Introduction]

> …The proposed BSTCM can establish a new paradigm of the human-machine interactions and secure wireless communications.

[System configuration]

> …As schematically demonstrated in Fig. 1b, a legal transmitter (Alice) equipped with an EEG cap intentionally transfers two visual secret ciphertexts to two legal receivers (Bob and Carol) at two harmonic frequencies based on the encrypted communication….

[Conclusions]

> …The proposed BSTCM, integrating the EM manipulation capability with the intelligence of the human brain, provides a new paradigm for the interactions among human-machine interfaces and the future 6G wireless communication applications.

Comment 3.3: *One of the biggest difficulties is the lack of focus on key features. Besides the very general overview provided in Figure 1, which is not self-explanatory, the article does not emphasize the crucial points. Instead, all arguments are presented at the same level, leaving the reader to discern the most important aspects. The authors seem to expect the reader to do the work of prioritizing the content.*

Response: We sincerely thank you for this valuable feedback. We understand the concern regarding the lack of focus on key features and the presentation of arguments at the same level, which may make it challenging for readers to identify the most crucial aspects of our work. To address this, we have made the following improvements. As stated in the response to ***Comment 1.1***, the key feature of this work is the innovative **fusion-coding method between SSVEP visual stimulation coding and metasurface space-time coding (STC) for secure information interaction at the physical layer**. SSVEP-based EEG frequency responses can be generated through visual stimulation frequency coding, while the STC metasurface controls the harmonic frequencies of EM waves. These frequency response characteristics exhibit certain similarities, forming a physical basis for their integration. Furthermore, the visual stimulation frequency is much lower than the switching frequency of the STC matrices, making it feasible to integrate visual stimulation coding for SSVEP-based BCI with STC into a fusing control signal, as schematically illustrated in Figure R8a. To implement this fusion coding scheme, we have designed a BSTCM platform that integrates LEDs with the STC metasurface, enabling simultaneous visual stimulation and EM wave regulation. The integrated platform enhances system compactness, mobility, and reliability by minimizing external connections. Moreover, the proposed fusion coding scheme unifies visual stimulation coding with STC-based EM wave manipulation, enabling real-time interaction and improved system efficiency. Based on the fusion coding method, we have realized a secure wireless communication system and remote smart device control.

Therefore, the crucial point is the security performance of the BSTCM system. The current BCI hardware devices have predominantly focused on enhancing their accuracy, functionality, and user-friendliness, but neglect the security of BCI data acquisition and transmission. Consequently, ensuring high security and privacy preservation becomes imperative when facilitating intelligent interactions within a complex communication environment. Based on the presented BSTCM system, we propose

a secure encrypted wireless communication system to realize reliable and secure information transmission between the human brain and external devices. The physical layer security of the proposed system can be realized by constructing different frequency-dependent communication channels via space- and frequency-multiplexing. The target information is encrypted by the proposed cryptography encryption method (refer to the response in ***Comment 3.1***). To verify the effectiveness of the proposed BSTCM system, we present its application in the remote control of smart devices. To emphasize this point, we have revised Fig. 1 and the related illustration is described in the section "System configuration".

Action Taken: Following your comment, we have supplemented a figure to demonstrate the core novelty in Figure 1. We have thoroughly revised the manuscript to enhance its structure, clarify the key features, and ensure that the most critical points are effectively emphasized. This adjustment enhances the clarity and readability of our manuscript. We hope that these revisions address the reviewer's concerns and improve the clarity, focus, and accessibility of our manuscript. We are grateful for the opportunity to refine our work and appreciate your valuable insights that have significantly enhanced the quality of our presentation.

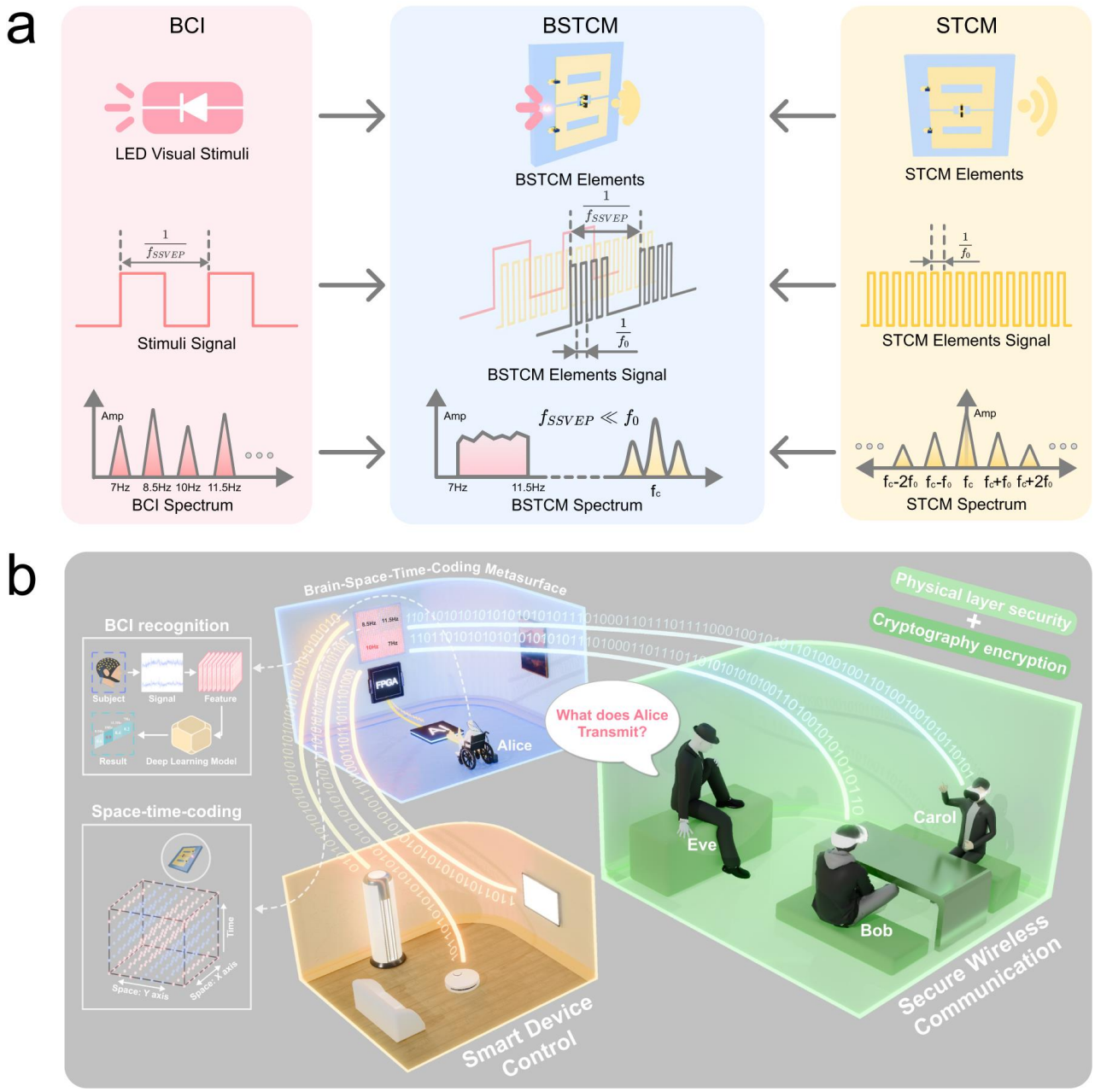

Figure R8. The schematical diagram of the presented BSTCM. a. The schematic diagram of fusing visual stimulation coding with the STC. The STC metasurface can correctly support the visual stimuli for the SSVEP-based BCI and facilitate information interaction with the external environment. b. The BSTCM system integrates STC metasurface into SSVEP-based BCI systems. Based on the BSTCM system, a secure encrypted wireless communication system is realized for the first time by combining with the variant HSKs and VSS method. In this way, smart devices can be controlled by the human mind with high security.

[Abstract]

…However, most existing BCI technologies are based on simple signal transmission and independent of other interface devices, with limited consideration for the reliability and security of human brain's information interaction in complicated wireless environments. To address the formidable limitation, we propose a deep fusion coding scheme that combines the BCI visual stimulation coding with metasurface space-time coding at the physical layer, enabling reliable and

secure information transfers between the human brain and external devices. To achieve this fusion coding scheme, we present a brain space-time-coding metasurface platform to implement a secure wireless communication system by using harmonic-encrypted beams. We design and fabricate a proof-of-principle prototype and experimentally show that the proposed wireless BCI scheme can establish a remote but safeguarded paradigm for human-machine interactions and intelligent metasurfaces, providing a potential direction in the future secure wireless communications.

[Introduction]

Here, we propose a deep fusion coding scheme that combines SSVEP visual stimulation coding with metasurface coding for secure information interactions at the physical layer. To implement this scheme, we present a brain space-time-coding metasurface (BSTCM) platform that seamlessly integrates the visual stimulation for SSVEP-based BCIs with the information interaction to external environment, significantly improving the system compactness and reliability. To enhance BCI security, we propose a secure encrypted wireless communication system that integrates the physical layer security with the cryptography encryption method. Specifically, the encryption method combines variant harmonic-based keys (HSKs) with visual secret sharing (VSS) technology. Experimental results demonstrate high bit error rate (BER) of nearly 50% for the eavesdroppers and the secrecy capacity of approximately 1.9 dB, validating the security of the encrypted wireless communication system. Finally, smart device control is presented by using the BSTCM system, realizing intelligent human-machine interactions. The proposed BSTCM can establish a new paradigm of the human-machine interactions and secure wireless communications.

Comment 3.4: *The article also uses many acronyms (for example, STC is introduced in the abstract but only partially defined), which further complicates understanding.*

Response and Action Taken: Thank you for raising this question. To avoid potential misunderstandings, we have revised the main text, enhancing the clarity of the manuscript.

[maintext]

brain space-time-coding metasurface (BSTCM)
space-time-coding (STC)
brain-computer interface (BCI)

Comment 3.5: *Moreover, several concepts are mixed together in a way that adds to the confusion. For instance, the use of visual excitation for encoding information is not well explained. It is unclear how this process works—does the person in front of the LED panel select different panels in sequence? Or is it merely a method to monitor which part of the panel is being observed?*

Response: Thank you for your insightful comment. The detailed flow diagram of the BSTCM system is illustrated in Figure R10. The system begins with flickering visual stimuli, designed to elicit specified SSVEP in the operator's brain. These EEG signals are then captured by BCI amplifier devices and processed using machine learning algorithms to extract and classify the operator's interaction intentions, which correspond to different flickering frequencies. Once the interaction commands are identified, they are transmitted to an FPGA. The FPGA serves as the central control module, integrating the visual stimulation signals with the corresponding STC matrices to generate a BSTCM signal. This signal drives the metasurface, enabling it to perform dual functions: providing the visual stimuli required for interaction and supporting advanced applications such as encrypted wireless communication systems or remote smart device control. This innovative approach enhances system efficiency and compactness by unifying visual stimulation, signal processing, and application control within a single platform. It represents a significant advancement in secure and intelligent human-computer interaction systems.

In a secure wireless communication application, the STC metasurface is divided into four partitions: 8.5Hz, 10Hz, 11.5Hz, and 7Hz. Two optimized STC matrices (M1 and M2) are employed to steer beams towards specific angles: -15° at the $+1^{st}$ harmonic frequency and +30° at the -1^{st} harmonic frequency, respectively. The ASK modulation scheme is used to realize wireless communication in our work. Specifically, the left panels of the metasurface, corresponding to frequencies f_1=8.5Hz and f_2=10Hz, are dedicated to transmitting digital symbols "0" and "1" to User1 at the position of θ = -15°. Similarly, the right panels, associated with the right panels f_3=11.5Hz and f_4=7Hz, are used to send digital symbols "0" and "1" to User 2 at the position of θ = +30°. Four BSTCM signal coding schemes are presented in Figure R9. The digital symbol "0" or "1" for User 1 is encoded as the repeated data frames "10001000" or "10101010…", respectively. Similarly, the digital symbol "0/1" to User 2 is encoded as repeated data frames "11001100…" or "11101110…", respectively. To transmit information, the operator interacts with the system through visual stimuli. For instance, when the operator gazes at the 8.5Hz region corresponding to the digital symbol "0" for User 1, the corresponding EEG signals are extracted and recognized to a control command. The command drives FPGA to fuse the visual stimuli signal of 8.5Hz and the STC matrices of M2 to form a BSTCM signal of Figure R10a. Then, the metasurface is controlled to simultaneously support 8.5Hz visual

stimuli and send the symbol “0” to User1 by the corresponding BSTCM signal. According to the binary bit streams of VSKs, the operator sends the corresponding information of VSKs in sequence by staring at different panels.

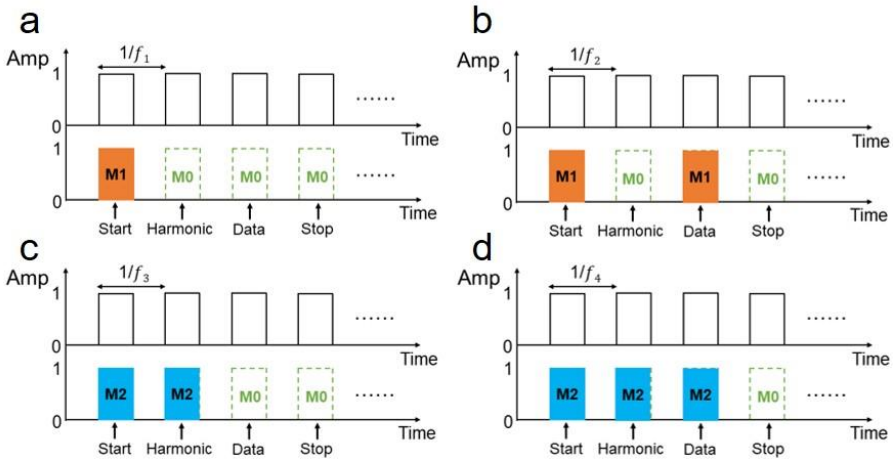

Figure R9. Four BSTCM signal coding schemes. a, b. The encoding schemes for the transmitting symbol “0” or “1” for two users: the symbols ‘“0” and “1” for Users 1. c, d. The encoding schemes for the transmitting symbol “0” or “1” for two users: the symbols ‘“0” and “1” for Users 2.

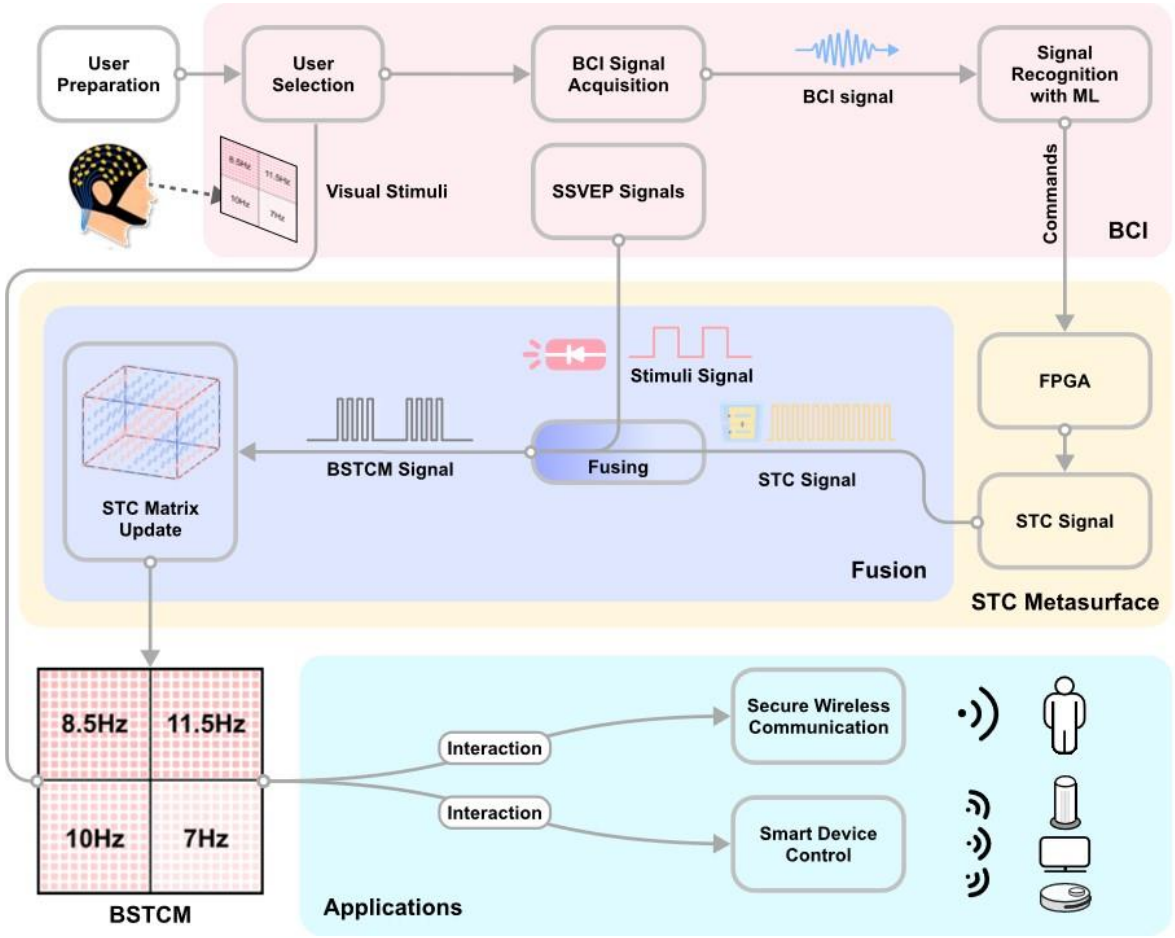

Figure R10. The detailed flow diagram of the BSTCM system. The user generates SSVEP signals by gazing at visual stimuli, which are processed and classified into different commands by the machine

learning method. These commands are sent to an FPGA, which fuses the visual stimuli signals and the STC signals to form different BSTCM signals for secure wireless communication and remote smart device control through frequency- and space-multiplexed physical channels.

Action Taken: We have added one section of "System configuration" to discuss the detailed flow diagram of the BSTCM system to help readers understand more clearly.

[System configuration]

> The schematic flow of the BSTCM system is presented in Fig. 2a. The EEG signals can be accurately recognized in the BCI system through the detection of SSVEP components. The target flickering stimuli with four frequencies are supported by the STC metasurface integrated with LEDs, and the commands can be output by simply fixating on the corresponding visual stimuli. The STC metasurface is composed of 32×32 elements, and segmented into four partitions, with each region consisting of 16×16 elements. In these partitions, LEDs are operated at four distinct frequencies (8.5Hz, 10Hz, 11.5Hz, and 7Hz) and implemented to evoke four distinct SSVEP signals. The interaction process begins with the user wearing an EEG cap with electrodes positioned over the O1 and O2 regions and sitting in front of the STC metasurface. Four visual flickering stimuli with distinct frequencies are presented to the user for selection. These flickering stimuli generate the SSVEP components in the user's BCI signals, while an EEG amplifier collects the EEG data (for details, see Supplementary Note 1). Following the sampling of BCI signals by the EEG cap, the BCI signals are later processed by a machine-learning algorithm, which classifies the signals into four different interaction commands, each corresponding to one of the aforementioned visual flickering frequencies. The interaction commands are then sent to FPGA, and the corresponding STC signals are generated. A fusion operation is conducted to fuse both STC signals and visual stimulation signals, hence generating the corresponding BSTCM signals. These BSTCM signals drive the STC metasurface to simultaneously maintain the visual stimuli of the fixed frequency and manipulate EM waves. Finally, the BSTCM system is used to construct independent physical channels at harmonic frequencies, enabling secure wireless communication and remote control of smart devices. This integrated platform not only ensures high mobility and system compactness by reducing the external connections but also achieves real-time bidirectional interaction. This approach highlights a novel fusion of BCI and STC metasurface technologies for secure, efficient, and intelligent human-computer interaction systems.

Comment 3.6: *Additionally, there seem to be as many LEDs as unit cells on the metasurface, yet these two effects appear to be completely decoupled.*

Response and Action Taken: Thank you for your comment. As illustrated in Figure R4a, we integrate visual stimulation coding for SSVEP-based BCI with space-time coding to effectively control the BSTCM, thereby enabling concurrent visual stimulation and information interaction. Hence, each unit cell integrates an LED, which shares the same voltage with the PIN diode. The visual stimulation of

BCI can be supported by the LEDs of the BSTCM element and the EM manipulation is realized by changing the PIN diodes' states of the element. All LEDs and PIN diodes loaded on the elements are simultaneously controlled by the same voltage signals with the help of an FPGA. We have made some revisions upon the section "Design of STC metasurface" in the revised manuscript.

[Design of STC metasurface]

> …To enable simultaneous visual stimulation and EM modulation in Fig. 4a, an LED serving as visual stimulation is placed outside the metal patch to avoid affecting the scattering properties of meta-atoms and share the same voltage signals with the PIN diode….

***Comment 3.7**: Why is space-time coding using a metasurface chosen? Instead of this complex setup, why not simply use a classical antenna or an array of antennas?*

Response: Thank you for your insightful comment. The decision to utilize a metasurface with STC instead of a classical antenna or antenna array is driven by the unique capabilities and advantages that metasurfaces provide for this specific application. The key benefits of employing metasurfaces for STC include the following:

1) Metasurface can integrate visual stimulation with EM wave modulation.

Metasurfaces offer a unique capability to integrate visual stimulation with EM wave modulation, which is challenging for traditional antenna arrays. In the proposed system, multiple LEDs used for visual stimulation are seamlessly embedded into the metasurface elements. These LEDs share the same voltage control signals as the PIN diodes within the metasurface, enabling integrated control through an FPGA. In contrast, traditional array antennas are inherently limited in their ability to perform such dual functionalities. Their design is primarily optimized for transmitting and receiving EM waves, without the structural or functional adaptability to incorporate visual stimulation. Moreover, the control mechanisms in array antennas are typically specialized for RF signal processing and do not support the simultaneous management of LEDs or other components. As a result, achieving the same degree of integration and efficiency as the metasurface-based system is practically difficult for conventional antenna arrays.

2) High integration and miniaturization

Metasurfaces are artificially engineered EM materials that are composed of specific geometric

structures arranged in periodic or quasi-periodic arrays, which can flexibly manipulate EM wave properties at the subwavelength scale. The thickness of metasurfaces can be much smaller than the wavelength, making the design and manufacturing more compact. Compared to antenna arrays, metasurfaces eliminate the need for bulky components like phase shifters and modulators, hence reducing system size, weight, and power consumption. Additionally, metasurfaces simplify the overall system architecture and lower production costs, making them a more efficient and economical choice for advanced communication systems.

3) Dynamic and flexible EM wave manipulation

Metasurfaces provide precise control over the amplitude, phase, polarization, and frequency of EM waves. By integrating active elements such as PIN diodes, varactors, liquid crystals, or other tunable materials, metasurfaces can dynamically manipulate the EM wavefronts in real time. This is achieved through FPGA-based control, enabling rapid and precise updates to the STC matrices. This capability is critical for STC, which requires real-time modulation of EM waves in both space and time domains by utilizing FPGA to quickly update STC matrices at an unprecedented level of precision and efficiency.

In summary, metasurfaces are the optimal choice for STC due to their precision, dynamic programmability, and high integration. Metasurface, as a compact and multifunctional platform, holds significant potential to deliver exceptional performance and scalability across a wide range of applications. The ability to integrate LEDs for visual stimulation directly into metasurface elements minimizes external connections and improves system reliability. This seamless integration facilitates unified control of both visual stimuli and EM wave manipulation, streamlining the system architecture.

Action Taken: We have revised the manuscript to clarify these points and better highlight the reasons behind our design choices. Thank you again for your constructive feedback.

[Introduction]

> …Wireless transmission of brain signals in the BCI systems is vulnerable to theft and attacks, potentially leading to inaccurate control commands and unauthorized privacy breaches. Although some methodologies have been proposed to enhance the security and privacy in BCIs23-25, the encryption mechanisms specifically tailored to these systems remain largely unexplored. Moreover, visual stimulation is often isolated from back-end information processing, lacking deep integration and interaction. With the increasing demand for secure BCI systems, it is essential to develop

intelligent interactions in secure and reliable communication environment. The frequency-dependent SSVEP responses and programmable harmonic characteristics of space-time-coding (STC) metasurfaces have notable similarities. Therefore, the STC metasurfaces can be used as a promising method that not only provides visual stimulation but also enhances the security of the BCI systems at the physical layer owing to their powerful ability to flexibly manipulate electromagnetic (EM) waves in both time and space domains26.
…Hence, the STC metasurface is a potential candidate for deep information modulation and interaction in the SSVEP-based BCI system owing to its capability to flexibly manipulate the EM waves and interact with the frequency-dependent visual stimulation.

Comment 3.8: *Why is the information transmitted to two users? Is the same information transmitted to both users? If so, why two users and not one or more?*

Response: Thank you for your professional comment. As stated in the response to ***Comment 3.1***, the target information is encrypted into two encrypted bit streams (VSK1 and VSK2) accompanied by two harmonic secret keys (HSK1 and HSK2) according to the proposed encryption encoding scheme. Two different encrypted bit streams are transmitted to two legal users by constructing two harmonic frequency-dependent physical channels, respectively. Each shared VSK remains devoid of any discernible information pertaining to the original target information, which effectively guarantees that the disclosure of any single VSK does not compromise the confidentiality of the original secret information. However, the eavesdroppers are relatively easy to intercept and decipher the target information when the target information is only sent to one user. In fact, the proposed encryption encoding scheme using two harmonic frequencies has great potential to be extended to include more harmonic frequencies in the encryption scheme, which further enhances the security of the system. For every additional user, an independent physical channel needs to be added, which puts higher demands on hardware and spectrum resources. As an illustrative example, we only present a dual-user encryption scheme, which provides higher security with less system complexity and a more efficient choice. In the future, we envision that more users are used to encrypt the communication system to achieve higher security.

Action Taken: Following your comment, we have added some explanations to the relevant sections of our manuscript. We appreciate your comment, which has helped us provide a clearer and more

comprehensive explanation.

[Secure encrypted wireless communication system]

> …Consequently, neither VSK1 nor VSK2 alone contains sufficient information to reveal the original secret information….
> …In the future, the encryption scheme can be expanded to include more users, further enhancing the BSTCM system's security.

***Comment 3.9**: The article also mentions that the pixel is split in two, which seems impossible.*

Response and Action Taken: Thank you for posing this question. We are so sorry that our previous statement caused you to misunderstand. As stated in the response to ***Comment 3.1***, the target information is encrypted into two encrypted bit streams accompanied by two pre-designed harmonic secret keys. The target information is composed of 5×5 pixels and each pixel is encoded into two encryption pixels according to the encryption coding scheme. We have changed the relevant expression in the manuscript.

[Secure encrypted wireless communication system]

> The conventional VSS method is employed to encrypt the individual pixels of confidential information, where each pixel is encrypted into two sub-pixels for two users.

***Comment 3.10**: Why are transmissions made over several harmonics? Isn't the focusing effect sufficient?*

Response: Thank you for your insightful comment. The encryption method is stated in response to ***Comment 3.1***. Based on the presented encryption method, the target information is encrypted into two VSKs accompanied with two HSKs. In the presented wireless communication system, two VSKs are sent to two receivers by constructing two different harmonic frequencies-based physical channels. In principle, it is feasible to construct two or more harmonic-frequency-dependent physical channels for secure wireless communication, allowing the target information to be divided into multiple parts. This significantly increases the difficulty of decryption. As a demonstration, we utilized two harmonic frequency users to validate the proposed encryption scheme. During the secure wireless

communication, the host user sends two VSKs sequentially to the corresponding users through the BSTCM platform. Firstly, the $+1^{st}$ order harmonic frequency physical channel pointing at -15° is constructed to transmit VSK1 to user 1. The physical channel is changed to the -1^{st} order harmonic frequency channel pointing at 30° to transmit the VSK2 to user 2 when the transmission of user 1 is completed. At any given moment, only one harmonic frequency channel is active, ensuring a strong focusing effect suitable for the proposed communication framework. Therefore, the focusing effect is sufficient for realizing the presented wireless communication. In future research, we can explore some more advanced algorithms to optimize STC coding strategies, which could enhance the performance of the focusing effect. Additionally, increasing the number of metasurface cells could further improve the system's focusing capabilities. These approaches offer promising pathways for further development and refinement of the technology.

Action Taken: Following your comment, we have added some discussion in the manuscript to enhance the quality of our work.

[Design of STC metasurface]

> The harmonic energy could be further enhanced in the future by optimizing STC coding strategies using advanced algorithms.

Comment 3.11: *The discussion about the keys is also unclear—are they known to the receivers?*

Response and Action Taken: Thank you for your good comment. As stated in response to ***Comment 3.1***, the harmonic secret keys are randomly generated according to the encryption scheme and each visual secret ciphertext is equipped with a harmonic secret key. The length of each HSKs is equivalent to that of the corresponding VSKs, and the HSKs dynamically adapt based on the content of the VSKs. These HSKs are shared with the legal receivers. We have added the discussion about keys in the manuscript.

[Secure encrypted wireless communication system]

> …The harmonic frequencies associated with the two shared sub-pixels are further integrated into the corresponding HSKs, which are securely distributed to two legal users. Notably, the length of each HSK is equivalent to that of the corresponding VSKs, and HSKs dynamically adapt based on

> the content of VSKs. This dynamic variation enhances the system's security by ensuring that the HSK remains intrinsically tied to the encrypted VSK….

Comment 3.12: *The second part of the article, which discusses control, is equally confusing. Is there any security feature involved, or is it merely a demonstration of remote control using EEG?*

Response and Action Taken: Thank you for your professional suggestion. The BSTCM system is employed to modulate EM wave characteristics across both spatial and frequency domains. This approach facilitates the construction of physical channels with distinct harmonic frequencies and beam-pointing directions, enabling the remote manipulation of smart user devices. In this work, four physical channels with harmonic frequencies (-2^{nd}, -1^{st}, $+1^{st}$, and $+2^{nd}$) corresponding to different harmonic pointing directions are constructed to remotely control four user devices, respectively. The properties of EM waves are utilized in this process to demonstrate the physical layer security assisted by the BSTCM platform, as stated in response to ***Comment 3.1***. We have discussed it in the manuscript. [Mind-controlled smart devices]

> …The excellent performance of harmonic beam manipulation demonstrates that the system can effectively utilize harmonic frequencies to construct four independent physical channels, significantly enhancing the physical layer security….

Comment 3.13: *In conclusion, it is very difficult to judge the effectiveness of the work because it is so confusing and mixes so many different concepts without providing the necessary keys to understand not only the choices made but also how the system works. For these reasons, I recommend rejecting the paper.*

Response: Thank you for your professional comment. We are so sorry for our confused expression in the original manuscript. We would like to address your confusion from two aspects and simultaneously point out the reasons for these choices:

complex wireless environments. Traditional setups typically use separate display screens or LEDs for BCI visual stimuli, which increases the risk of secret information leakage. **To address this problem, we propose the deep fusion of SSVEP visual stimulation coding with metasurface STC in one integrated platform for secure information interaction, as shown in Figure R10.** The presented fusion coding scheme enables real-time interaction and improves system efficiency. To realize the fusion coding scheme, we have designed an integrated platform by embedding LEDs onto the STCM, thereby simultaneously supporting visual stimulation and information interaction. The proposed integrated platform reduces the risk of connection failures by reducing external connections and physical interfaces, thereby improving the overall reliability of the system. Furthermore, the frequency-dependent SSVEP response and the programmable harmonic characteristics of STC metasurfaces exhibit inherent compatibility, making their integration not only feasible but also highly effective. The seamless embedding of LEDs into the metasurface further highlights the practicality of this design. Therefore, the STC metasurface provides an ideal solution for achieving secure information interaction at the physical layer.

Secondly, <u>for the system working process,</u> as depicted in Figure R10, the schematic overview of the procedures involved in our BSTCM system is presented. The process begins with the user wearing an EEG cap and sitting in front of the STCM. Four LED visual flickering stimuli, each with a distinct frequency, are displayed for the user to select.

1) Once the BCI signals are sampled, they undergo processing and recognition through a **machine-learning algorithm**, as detailed in Figure R11. First, BCI signals are pre-processed to generate SigSpec through decomposition, spectrum calculation, and weighted summation. Second, SigSpec undergoes an outer product operation with reference spectra RefSpec, producing eight feature graphs that highlight a square grid pattern at the target frequency. **These feature graphs are then classified using a lightweight CNN**, enabling accurate classification of BCI signals into four interaction commands. The recognized command is sent to an FPGA, which performs a fusion operation to combine the STC metasurface signals and visual stimulation signals, generating BSTCM signals. These signals drive the STC metasurface to simultaneously sustain visual stimuli of fixed frequencies and manipulate EM waves for various tasks.
2) Based on the fusion coding strategy, a secure wireless communication system is implemented to enable secure information transfer between the human brain and external devices. As outlined in

response to ***Comment 3.1***, the proposed wireless communication system integrates **physical layer security** with a **cryptographic encryption method**, enhancing its security performance. Embedding security directly into the frequency-dependent physical channels makes the encryption process more resilient to attacks, such as brute force decryption or algorithmic vulnerabilities. **Furthermore, we propose a cryptography encryption method that combines the VSS method with variant secret keys, where the harmonic frequencies of the BSTCM serve as the secret keys.** The target information is firstly encrypted to two encrypted bit streams accompanied by two HSKs based on the present encryption method. Then, two harmonic frequency-dependent physical channels are constructed to send the relevant VSKs to two specified users by the BSTCM platform.

3) The remote smart device control is then demonstrated to validate the effectiveness of the BSTCM platform. **Four harmonic frequency-dependent channels, each directed toward a distinct spatial angle, are constructed using the four partitions of the BSTCM platform, thereby ensuring physical layer security.** Each partition points in a different direction by updating the optimized STC matrix. To control a device, the user simply gazes at the corresponding partition of the BSTCM, enabling direct and remote activation of the target device.

We are committed to addressing these points in the revised version to strengthen our findings. We sincerely hope this clarification will help the referee gain a better understanding of our work and reconsider their decision.

Action Taken: To clearly demonstrate our work, we have revised the corresponding figures and discussion in the manuscript. We thank **R3** for the comment, which is very helpful to enhance the quality of our manuscript.

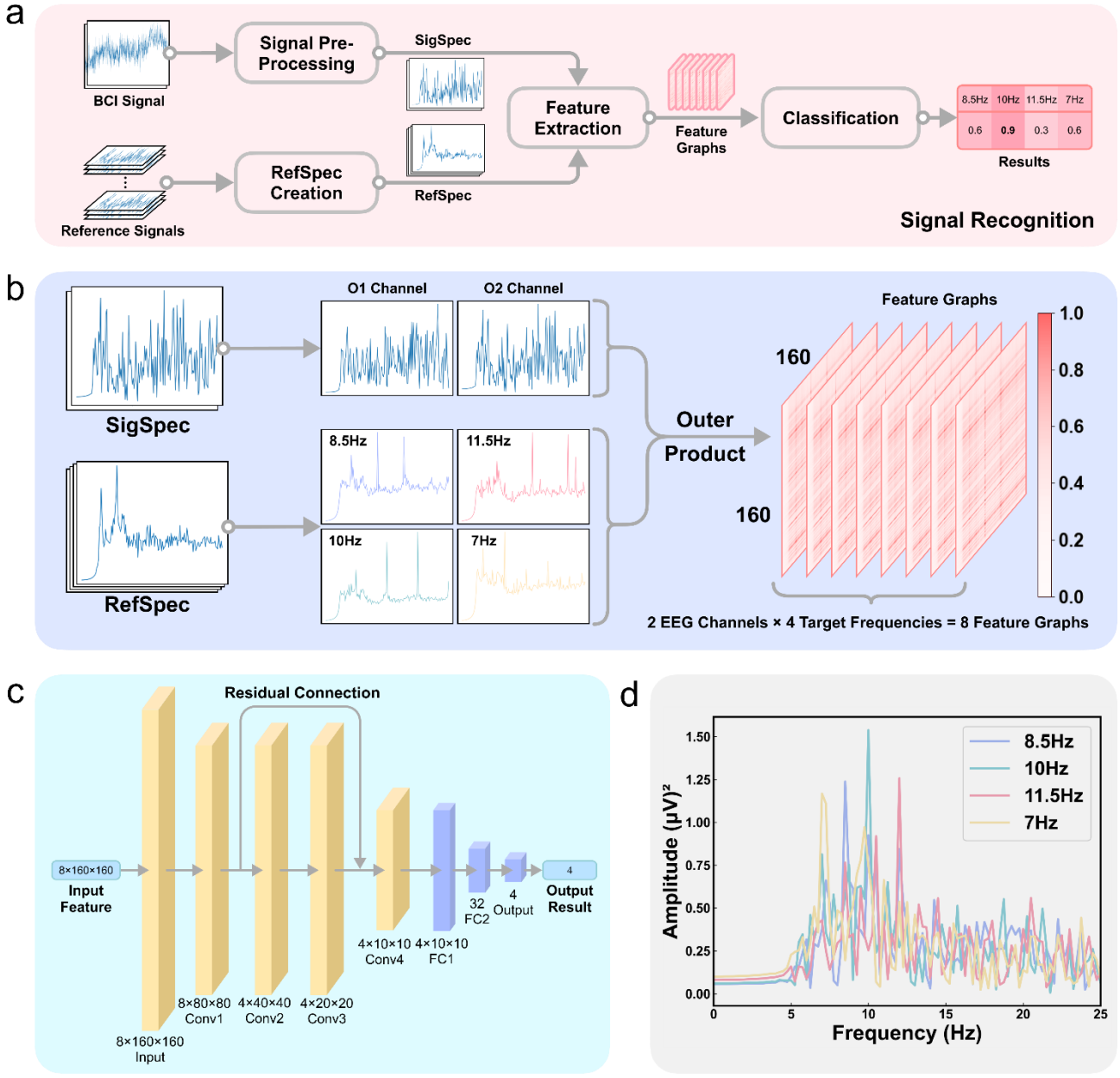

Figure R11. The detailed diagrams of signal recognition of the SSVEP-BCI system. a. The comprehensive workflow from raw BCI signal acquisition to the final signal recognition outcome. **b.** Details the conversion of SigSpec into eight distinct feature graphs through an outer product operation with pre-calculated RefSpec. **c.** The lightweight classification model. Four convolutional and two fully connected layers were employed. **d.** An example of frequency responses of SSVEP signals for the four target frequencies, showing distinctive amplitude peaks corresponding to different frequencies.

[Abstract]

> …However, most existing BCI technologies are based on simple signal transmission and independent of other interface devices, with limited consideration for the reliability and security of

human brain's information interaction in complicated wireless environments. To address the formidable limitation, we propose a deep fusion coding scheme that combines the BCI visual stimulation coding with metasurface space-time coding at the physical layer, enabling reliable and secure information transfers between the human brain and external devices. To achieve this fusion coding scheme, we present a brain space-time-coding metasurface platform to implement a secure wireless communication system by using harmonic-encrypted beams. We design and fabricate a proof-of-principle prototype and experimentally show that the proposed wireless BCI scheme can establish a remote but safeguarded paradigm for human-machine interactions and intelligent metasurfaces, providing a potential direction in the future secure wireless communications.

[SSVEP signal recognition]

In this study, we present a robust deep learning-based algorithm for the recognition of SSVEP component, which is seamlessly integrated into the signal recognition stage of the proposed system (refer to Supplementary Note 2). As depicted in Fig. 3a, the raw BCI signals undergo a comprehensive pre-processing pipeline to extract the key frequency features for SSVEP signal recognition. This pipeline filters the signals using a filter bank consisting of four Chebyshev Type I bandpass filters, isolating four distinct sub-bands (SB1-SB4) with specified frequency ranges to mitigate the interference and enhance the signal quality. Subsequent application of fast Fourier transform (FFT) to these sub-bands reveals their amplitude spectra, highlighting the frequency components of the SSVEP signals. To further refine the spectral information, a weighted summation process is used to assign weight factors to each sub-band based on their indexes, emphasizing the contribution of lower frequency components. The resulting signal amplitude spectrum (SigSpec) encapsulates the essential frequency characteristics of the SSVEP signals, serving as the foundation for subsequent feature extraction and classification processes.

Furthermore, we employ an innovative feature extraction technique to perform outer product operations between the SigSpec and pre-computed reference spectra (RefSpec), derived from the SSVEP signals with known frequencies, as described in Fig. 3b. This operation yields eight feature graphs (sized by 160×160 pixels), each representing the interaction between two input SSVEP signal channels (O1 and O2) and the reference signals corresponding to the four target frequencies. A discernible square grid pattern appears in the feature graph corresponding to the frequency of the anticipated classification result (refer to Supplementary Notes 3-5). These feature graphs are then processed by a lightweight CNN-based classification model, which incorporates four convolutional layers and two linear layers, as shown in Fig. 3c and Supplementary Figure 12. The feature graphs are initially processed through a convolutional input layer. Subsequently, a MaxPooling operation with a 2×2 kernel is applied after the data are activated by the ReLU function. To facilitate the propagation of original data, a residual connection is strategically integrated between the second and third convolutional layers. Each convolutional layer is followed by the same 2×2 MaxPooling operation and ReLU activation. Once the convolutional layers extract spatial feature vectors from the feature graphs, a series of two consecutive linear layers is employed to classify the feature vectors into four distinct target frequencies. Finally, an output layer with sigmoid activation reshapes the model's output into four confidence scores. After training 300 epochs, the model achieved a classification accuracy of 96.67% on the validation set, with precision and recall rates over 90%. These results highlight the effectiveness of the proposed deep learning-based algorithm

for the SSVEP signal recognition (see Supplementary Notes 6 for additional details).

Comment 3.14: *Line 77: "Hence, the STC metasurface is a … ." I do not understand the sentence. Is the STC processing the information? How?*

Response and Action Taken: We appreciate the **R3**'s valuable question. The STC metasurface is used to dynamically modulate the properties of EM waves, such as phase, amplitude, or frequency, which can be used to encode and transmit information. The STC metasurface does not directly process information. Our expression may not be very accurate. We have revised the manuscript. We thank **R3** for the important comment, which is very helpful in enhancing the quality of our manuscript.

Action Taken:

[Introduction]

…Hence, the STC metasurface is a potential candidate for deep information modulation and interaction in the SSVEP-based BCI system owing to its capability to flexibly manipulate the EM waves and interact with the frequency-dependent visual stimulation.

Comment 3.15: *Line 84: "Information interaction" – What does this mean?*

Response: Thank you for raising the question. In this context, **"information interaction"** refers to the bidirectional exchange of information between the BSTCM system and the external environment. The visual stimulation generated by the STCM elicits specific SSVEP in the user's brain, which are detected and recognized as commands. The BSTCM system translates the user's brain commands into specific instructions that are used to control external devices or facilitate secure wireless communication. The BSTCM system enables the user to send commands by the brain and facilitating system responses to execute tasks in the external environment. In the presented encrypted wireless communication, the target information is remotely transmitted to other users by the BSTCM platform. In addition, different commands are sent to control smart devices by the host user. The term highlights the seamless integration of visual stimulation coding and space-time coding to achieve dynamic and intelligent interactions with the external environment.

***Comment 3.16**: Line 86: "High-security" – How high? Is it equivalent to 256-bit AES encryption? (I don't think so.)*

Response: Thank you for your professional comments. As addressed in response to ***Comment 3.1***, the proposed encrypted wireless communication system combines **physical layer security** and a **cryptographic encryption method**, significantly enhancing overall security. The target information is encrypted into two parts and transmitted to two legal users through distinct harmonic frequency-based physical channels. The information cannot be decrypted from a single channel alone, thus greatly improving security. To demonstrate the security, we conducted a comprehensive evaluation of the system's security performance. Figures R12a and b illustrate the system's robustness against eavesdropping attempts in both spatial and frequency domains, respectively. The results, summarized in Figures R12c and d, show that the measured **bit error rate (BER)** for illegal users remains consistently around 50%. Specifically, for an illegal user attempting to intercept the -1st harmonic frequency channel at various spatial positions (35° to 60°), the BER remains nearly 50%. Similarly, when illegal users with harmonic frequencies ranging from -2nd to +2nd order are tested at the 35° spatial position, the corresponding BERs are nearly 50%. Furthermore, the **signal-to-noise ratio (SNR)** for legal users is measured as approximately 22 dB. In contrast, the SNR for illegal users located in the 35° spatial direction is less than 5 dB. Hence, the **secrecy capacity** ($C_s = \log_2(1+S_{legal}/N_{legal}) - \log_2(1+S_{illegal}/N_{illegal})$) between the legal user and the illegal user is calculated as about 1.9 dB, demonstrating the high level of security provided by the BSTCM system. While the security mechanism does not directly compare to 256-bit AES encryption, the combination of physical layer security and cryptography encryption method, along with the highly secure design of the harmonic frequency-based channels, ensures robust protection against eavesdropping and unauthorized access.

Action Taken: Following your comment, we have added the security performance discussions in the main text and supplementary materials of the revised manuscript. We sincerely appreciate **R3** for raising such a valuable question, as it has greatly contributed to enhancing the quality of our work.

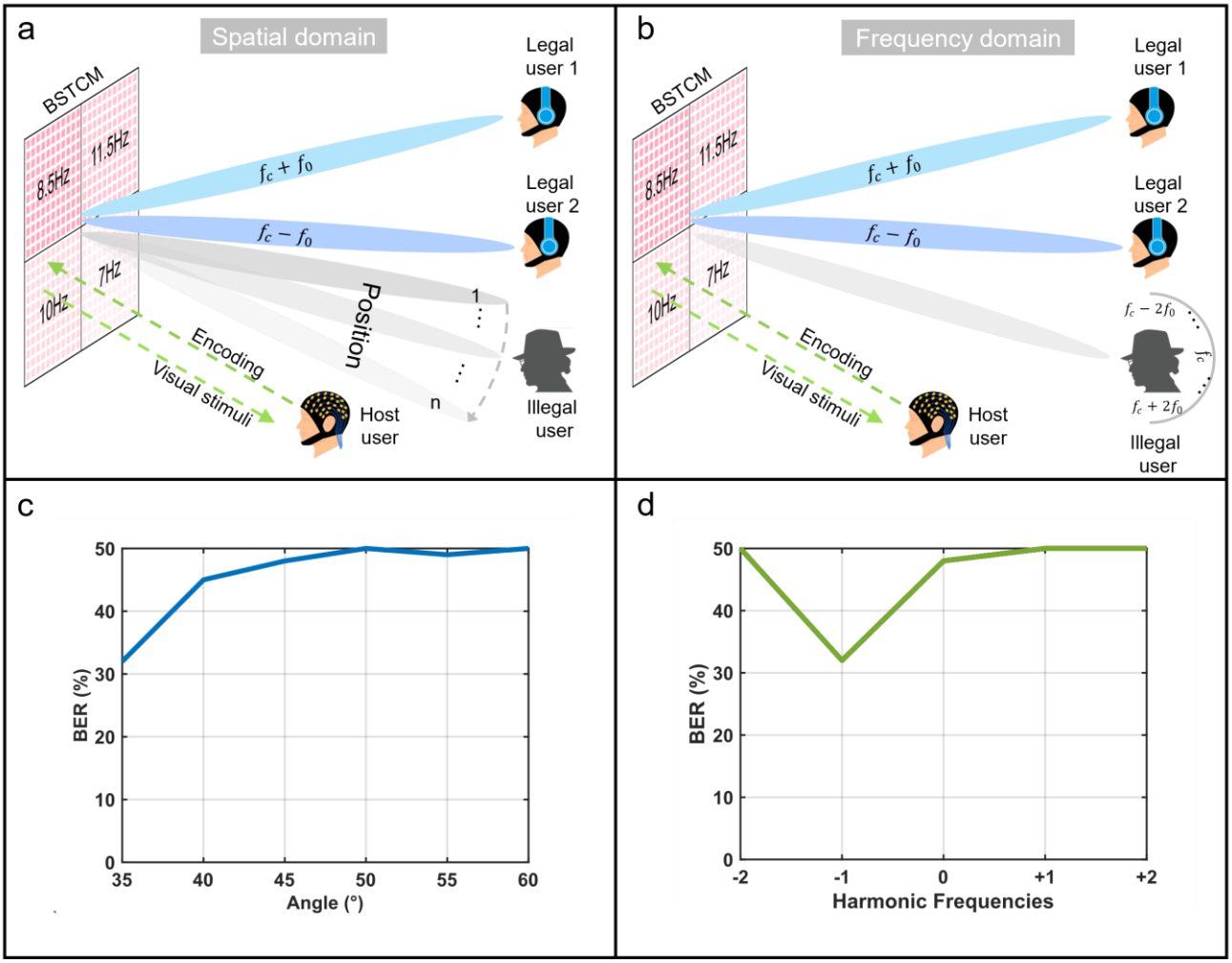

Figure R12. The security measurement of the proposed secure wireless communication scheme. a. The diagram of security test scenarios for illegal users in different spatial locations from 35° to 60°. b. The diagram of security test scenarios for illegal users in different harmonic frequencies from -2nd to +2nd order. c. The measured results for illegal users in different spatial locations from 35° to 60°. d. The measured results for illegal users in different harmonic frequencies from -2nd to +2nd order.

[Supplementary Note 15: The Security Performance of the Proposed Secure Wireless Communication]

We have presented a cryptography encryption method as described above, which combines the visual secret sharing (VSS) method and the variant harmonic-based secret keys. The proposed scheme distributes confidential information between two users, ensuring that interception of a single communication channel does not disclose any information of the shared secret. In BSTCM-assisted wireless communication systems, multiple harmonic-frequency-dependent physical channels can be constructed to realize multichannel information transmission because the BSTCM can control the propagation direction and harmonic power distribution of the EM waves in both the spatial and frequency domains, which offers the physical layer security that leverages the characteristics of the physical channel to achieve information encryption. To verify the security of the proposed scheme, we have tested the SNR and BER of the BSTCM system. As shown in Supplementary Figure 26a and b, the system security is measured when the illegal user eavesdrops separately from the spatial and frequency domains, respectively. Firstly, we measured the BER

results of the illegal user with -1st harmonic frequency in different spatial locations from 35° to 60° in the BSTCM-based wireless communication system. The BERs are nearly 50% shown in Supplementary Figure 26c. The illegal users with different harmonic frequencies from -2nd to +2nd order have been tested when they are located at 35° spatial direction. Supplementary Figure 26c demonstrated that BERs are nearly 50%. In addition, the signal-to-noise ratio (SNR) for the two legal users is measured as approximately 22 dB while the illegal user is less than 5 dB. The secrecy capacity is denoted as the difference in the channel capacity between the legal user and the illegal user. It is described as follows: $C_s = log_2(1 + S_{illegal}/N_{legal}) - log_2(1 + S_{illegal}/N_{illegal})$, where S_{legal} and $S_{illegal}$ are the powers received by the legal user and the illegal user and N_{legal} and $N_{illegal}$ denote the corresponding noise. The secrecy capacity of our system is calculated as about 1.9 dB. Consequently, the proposed scheme is highly secure and not easily cracked unless he/she can crack the channel direction, harmonic frequency, and two HSKs simultaneously.

Comment 3.17: *Line 100: "Regulate the EM" – EM produced by what?*

Response: Thank you for your comment. A commercial signal generator is capable of producing EM waves across a range of frequencies and power levels. These EM waves are then focused into a directional beam through a horn antenna, which is connected to the signal generator. Subsequently, the incident EM waves are modulated by the STC metasurface, thereby enabling the realization of various EM functions.

Comment 3.18: *Line 110: "The BSTCM system empowers individuals." What does this mean?*

Response and Action Taken: Thank you for your good comment. This statement refers to the unique capability of the BSTCM system, enabling users to control and modulate EM waves through their brain signals. Specifically, EEG signals are captured, which are then processed to generate commands. These commands are translated into instructions that control the STC metasurface, allowing it to dynamically manipulate EM waves such as beam direction and harmonic frequency generation. The BSTCM system acts as a bridge between the human mind and physical EM wave manipulation, empowering users to interact with and control the external devices by their brain signals. This integration showcases a novel human-machine interaction paradigm that combines metasurface technology and BCI for advanced control.

[System configuration]

> The BSTCM system empowers individuals to customize and modulate the EM waves by the human mind, enabling a variety of EM functions and enhancing the flexibility and interactivity of human-machine systems.

Comment 3.19: *Line 111: "Abundant harmonic…" – This needs clarification.*

Response and Action Taken: Thank you for your good suggestion. The STC metasurface can control the propagation direction and harmonic power distribution of the EM waves in both the spatial and frequency domains. The STCM is capable of generating and manipulating multiple harmonic frequencies of EM waves. This feature allows the construction of distinct physical channels, each operating at a different harmonic frequency and beam-pointing direction, enabling the system to separate and securely transmit information or control signals. The term emphasizes the richness of available frequencies and their role in expanding the functionality of the system. We have supplemented and revised in the manuscript.

[System configuration]

> The STC metasurface enables precise control to the propagation direction and harmonic power distribution of EM waves across both spatial and frequency domains. By harnessing the abundant harmonic modulation capabilities and integrating the advantages of human brain intelligence, the BSTCM system facilitates the implementation of a secure encrypted wireless communication system and enables smart device control by human mind….

Comment 3.20: *Line 115: "Transmitted information" – But which information? Sent by the user, or by the BSTCM?*

Line 116: "Visual secret ciphertexts" – Are these chosen by the user or by the BSTCM? If the latter, what is the interest?

Response and Action Taken: Thank you for raising these questions. The transmitted information refers to the original target information. As previously discussed in response to ***Comment 3.1***, the target information is firstly encrypted into two visual secret ciphertexts (VSKs) according to the presented encryption method. Then, two VSKs are transmitted to different users by the host user via

the BSTCM system. The visual secret ciphertexts are selected by the host user and transmitted to two legal users by the BSTCM system. The BSTCM system combines physical layer security with a cryptography encryption method, ensuring a secure and reliable communication process. We have revised our manuscript.

[System configuration]

> …In the encrypted wireless communication system, the harmonic frequencies are adopted as the secret keys and the VSS method is employed to encrypt the target information[46-48]. The target information is encrypted into two visual secret ciphertexts accompanied by two different HSKs….

Comment 3.21: *Line 122: I do not understand the Metaverse applications.*

Response and Action Taken: Thank you for raising the question. Numerous studies have shown that BCI technology is highly compatible with metaverse applications. Leveraging these findings, we propose a hypothesis that our BSTCM system could enhance and contribute to this field in upcoming advancements. In order not to confuse readers, we have removed the relevant statements. The work focus on the innovative fusion of visual stimulation coding for SSVEP-based BCI with space-time coding to control the BSTCM for simultaneous visual stimulation and information interaction, as previously discussed in response to ***Comment 3.3***.

[Abstract]

> …We design and fabricate a proof-of-principle prototype and experimentally show that the proposed wireless BCI scheme can establish a remote but safeguarded paradigm for human-machine interactions and intelligent metasurfaces, providing a potential direction in the future secure wireless communications.

[Introduction]

> …The proposed BSTCM can establish a new paradigm of the human-machine interactions and secure wireless communications.

[Conclusions]

> …The proposed BSTCM, integrating the EM manipulation capability with the intelligence of the human brain, provides a new paradigm for the interactions among human-machine interfaces and the future 6G wireless communication applications.

Comment 3.22: *Line 124: "There are now secret keys?" Are these different from the ciphertexts?*

Response and Action Taken: Thank you for your professional comment. As previously discussed in response to ***Comment 3.1***, the target information is encrypted into two visual secret ciphertexts accompanied with two harmonic secret keys according to the presented encryption method. Two harmonic secret keys are different with two visual secret ciphertexts. The encrypted information cannot be decrypted without two corresponding secret keys even if the eavesdropper eavesdrops on two ciphertexts. We have supplemented some descriptions in the manuscript.

[System configuration]

> …In the encrypted wireless communication system, the harmonic frequencies are adopted as the secret keys and the VSS method is employed to encrypt the target information[46-48]. The target information is encrypted into two visual secret ciphertexts accompanied by two different HSKs. As schematically demonstrated in Fig. 1b, a legal transmitter (Alice) equipped with an EEG cap intentionally transfers two visual secret ciphertexts to two legal receivers (Bob and Carol) at two harmonic frequencies based on the encrypted communication.

Comment 3.23: *Line 131-132: There is repetition here.*

Response and Action Taken: Thank you for your good comment. We have made some revisions.

[System configuration]

> …The proposed encrypted wireless communication system integrates the physical layer security with the cryptography encryption method, significantly enhancing the system security. Furthermore, the proposed system also facilitates remote control of smart devices, enabling users to control devices directly through the user's brain intention without the need for physical actions.

Comment 3.24: *Line 133: The spectrum is not completely clear. What is the difference between notes 2 and 5—recognition or extraction?*

Response: Thank you for your insightful comment. A clearer explanation is described as follows:

The signal spectrum (SigSpec) refers to the amplitude spectrum of newly collected brain signals, while the reference spectra (RefSpec) are the average amplitude spectra of reference signal datasets.

Regarding supplementary notes 2 and 5:

- Supplementary Note 2: This focuses on “SSVEP signals recognition”, detailing the entire process of using SSVEP-BCI in the study.
- Supplementary Note 5: This provides a detailed description of the feature extraction process, which is part of the recognition process. Feature extraction transforms raw signal data into feature graphs.

In summary, feature extraction and classification are two distinct stages. The extraction stage converts raw data into feature graphs, and the classification stage classifies these graphs into target classes.

Comment 3.25: *Line 173: "Such results underscore the efficacy of the SSVEP component…." I do not understand—was ML used or not?*

Response and Action Taken: Thank you for your professional comment. A machine learning algorithm was indeed utilized in our study to process and classify the SSVEP-based EEG signals. Specifically, the algorithms were employed to identify distinct SSVEP responses corresponding to the visual stimuli frequency, which allowed for the accurate translation of EEG signals into interaction commands. The statement emphasizes the successful EEG signals classification with the presented machine learning algorithm. We have revised the statement to make this connection clearer.

[SSVEP signal recognition]

> …These results highlight the effectiveness of the proposed deep learning-based algorithm for the SSVEP signal recognition (see Supplementary Notes 6 for additional details).

Comment 3.26: *Line 186: Why are the LEDs connected to the PIN diodes? Would it not have been simpler to decouple the two?*

Response and Action Taken: Thank you for your good comment. As previously discussed in response to ***Comment 3.6***, each unit cell integrates an LED, which shares the same voltage with the PIN diode. We have successfully integrated visual stimulation coding for SSVEP-based BCIs with space-time coding to efficiently control the BSTCM. This integration facilitates simultaneous visual stimulation and information interaction. All LEDs and PIN diodes mounted on the elements are concurrently controlled by the same voltage signals, managed via an FPGA. Additionally, we have made several

revisions to the section titled “Design of STC Metasurface” in the revised manuscript to reflect these advancements. We have made some revisions upon the section “Design of STC metasurface” in the revised manuscript.

[Design of STC metasurface]

> …To enable simultaneous visual stimulation and EM modulation in Fig. 4a, an LED serving as visual stimulation is placed outside the metal patch to avoid affecting the scattering properties of meta-atoms and share the same voltage signals with the PIN diode….

Comment 3.27: *Line 246: The article speaks of images without explaining that images are being transmitted.*

Response and Action Taken: Thank you for raising the question. We apologize for any confusion caused by our earlier statement. The images referenced in the manuscript are binary, consisting solely of black and white pixels, encoded as digital “0” and “1”. Specifically, the “H” image depicted in the figure is a binary matrix of 0 and 1, structured as a 5×5 pixel grid. We have revised the relevant section in the manuscript to clarify this point. We appreciate **R3** insightful comments, which have significantly contributed to improving the quality of our manuscript.

[Secure encrypted wireless communication system]

> …It is worth noting that each share derived from this process remains devoid of any discernible information pertaining to the original target information….
> …As an illustrative example in Fig. 5a, the secret image “H”, consisting of a grid of 5×5 black pixels (digital “1”) and white pixels (digital “0”), is encrypted into two random VSK1 and VSK2 based on the presented encryption method….

Comment 3.28: *Lines 254-305: I do not understand where the brain coding occurs.*

Response: Thank you for raising the question. As stated in response to ***Comment 3.1*** and ***Comment 3.5***, the target information is firstly encrypted into two VSKs. When the operator focuses on a defined flickering region representing the content of the VSKs, the corresponding EEG signals are extracted and translated into a command. This command directs the FPGA to fuse the visual stimuli signal with the STC matrices, generating a BSTCM signal. The metasurface is then controlled to simultaneously

provide the corresponding visual stimuli and transmit the VSK to the user via the BSTCM system. Based on the binary bit streams of the VSKs, the operator sequentially transmits the corresponding VSK information by gazing at different panels.

Authors' Responses

Firstly, we appreciate all reviewers for the constructive and insightful comments and suggestions, which helped us improve the quality of our manuscript significantly. Based on these comments and suggestions, we revised the manuscript very carefully. Below are our point-by-point responses to the reviewers' comments.

REVIEWER #1

We would like to thank this reviewer for the positive comment regarding to our manuscript. In particular, the reviewer states that "*While the revision has addressed some of the previous questions….*", and provides several suggestions on how to improve our manuscript. We thank the reviewer for the positive feedback on our manuscript and do appreciate the suggestion. The responses to the suggestions and the changes we have made in response to the comments are outlined below.

Comment 1.1: *While the revision has addressed some of the previous questions, it remains unclear how information is transmitted from the EEG signal to other users. Additionally, the data transmission process needs further clarity, particularly regarding the amount of data that can be transmitted in terms of bits.*

Response: Thank you very much for your insightful comments. We fully understand your confusion regarding to the perceived complexity of the proposed system. To put it more simply, you can think of the metasurface as a visual keyboard, and the user selects information by looking at the keys (the four areas on the metasurface on this "keyboard"). When the user selects this key (by continuous gaze), BSCTM can directly obtain the user's selected key through the user's BCI signal analysis and modulate the corresponding information onto the metasurface. The metasurface with the STC modulation will radiate toward the target user in a specific direction under the excitation of the horn antenna, completing the information transmission process. In the information decoding process, the two users at the receiving end can obtain the information after receiving the corresponding encoded information in combination with the known decoding rules.

To clearly illustrate the data transmission process, we clarified the secure wireless communication process of our work. According to the proposed encryption schemes of Fig. 5a, the transmitted information is firstly encrypted into two binary ciphertexts (VSK1 and VSK2). The system begins with visual stimuli, designed to elicit specified SSVEP in the operator's brain. The detailed flow diagram of the BSTCM system is illustrated in Figure R1. When the host user intends to transmit the message to other users, signal transmission can be achieved by gazing at the metasurface.

The specific process is as follows:

Step 1: In the proposed secure wireless communication process, the host user can subjectively choose the information they want to send by gazing at specific areas of the metasurface. Specifically, the host user can send the binary symbol "0" and "1" of VSK1 to User 1 by gazing at the left panels (f_1=8.5Hz and f_2=10Hz) of the STC metasurface. By staring at at right panels (f_1=11.5Hz and f_2=7Hz), the binary symbol "0" and "1" of VSK2 can be sent to User 2. The EEG signals of four visual flicker stimulation frequencies are different. Then, the corresponding EEG signals are then captured by BCI amplifier devices and processed using machine learning algorithms to extract and classify the host user's interaction intentions, which correspond to different flickering frequencies. The interaction intentions corresponding to different symbols of VSKs are classified into different interaction commands.

Step 2: Different interaction commands are then transmitted to an FPGA, which integrates the visual stimulation signals with the corresponding STC matrices to generate a BSTCM signal. As illustrated in Figures 4j and 4k, the fusion process of a BSTCM signal follows a structured encoding scheme based on an amplitude-shift keying (ASK) modulation scheme, where the symbol "0" or "1" of VSK1 is represented by periodically repeated data frames: "M1-M0-M0-M0-M1-M0-M0-M0…" or "M1-M0-M1-M0-M1-M0-M1-M0…". The fused BSTCM signal cycles periodically, with each data frame consisting of four bits, where the third bit of each frame serves as the transmitted data bit and the remaining bits are used as check bits. Similarly, the digital symbol "0" or "1" to User 2 is encoded as repeated data frames "M2-M2-M0-M0-M2-M2-M0-M0…" or "M2-M2-M0-M0-M2-M2-M0-M0…"., respectively. The BSTCM signals are then

used to drive the STC metasurface, which generates directional beams at specific harmonic frequencies and directs them toward two users for secure wireless communication.

In our system, the bit rate is primarily constrained by the performance limitations of the BCI. Specifically, due to the inherent characteristics of SSVEP and the signal processing requirements for accurate EEG decoding, we employ a 2ASK modulation scheme. Given the need for reliable classification of brain responses, each symbol is transmitted over a 4-second window, resulting in a bit rate of 0.25 bps. While this bit rate may appear low compared to conventional wireless communication systems, it is important to highlight that BCI-based communication systems have lower transmission rates due to neural response time constraints and the necessity for high classification accuracy. In this case, our system prioritizes robustness and security over high-speed transmission. Additionally, our proposed BSTCM-based secure wireless communication framework ensures that even with a low bit rate, critical information such as control commands, authentication data, or encrypted messages can be effectively transmitted.

It is worth mentioning that this bit rate may be further improved with the development of technology, mainly in the following three aspects: (1) more precise and sensitive BCI electrodes; (2) smarter and more efficient BCI signal processing algorithms; (3) higher metasurface control rate, which mainly depends on higher performance transistors.

1. More Precise and Sensitive BCI Electrodes

The evolution of EEG-based BCI systems hinges critically on advancements in electrode technology to achieve higher precision and sensitivity[R1]. Modern wireless EEG devices now support high-density configurations (up to 128 channels) with sampling rates exceeding 2 kHz, enabling ambulatory applications without the constraints of traditional wet electrodes. Furthermore, adaptive electrode arrays with embedded signal conditioning circuits are being explored to dynamically compensate for environmental noise and physiological artifacts, particularly in real-world scenarios involving head movement or muscle activity.

2. Smarter and More Efficient Signal Processing Algorithms

The advancement of EEG-based SSVEP-BCI technology is poised for significant acceleration through

the development of smarter and more efficient signal processing algorithms, as evidenced by recent methodological innovations. The multiscale octave convolution neural network concept, initially proposed for ERP detection[R2], presents a promising framework for SSVEP applications. The integration of filter bank techniques with deep neural networks, as demonstrated in recent studies achieving 94.29% accuracy through optimized frequency-band decomposition[R3], suggests further improvements in signal-to-noise ratio management. Moreover, the emergence of lightweight convolutional architectures like EEGNet[R4] and Compact-CNN[R5] indicates a pathway toward real-time processing optimization. Hybrid paradigm integration, such as the SSVEP-EOG combination achieving 117.05 bits/min information transfer rate through time-space frequency conversion[R6], exemplifies how multimodal signal fusion could enhance system robustness.

3. Higher Metasurface Control Rate

The advancement of metasurface control rates critically depends on transistor performance. The control rate of the metasurface can be significantly improved by employing transistors with higher switching speeds, thereby reducing the response time of individual unit cells. Furthermore, the integration of high-speed FPGA architectures significantly enhances the metasurface's control rate, thereby enabling ultra-fast reconfiguration of electromagnetic properties.

References:

[R1] Edelman, B. J. et al. Non-Invasive Brain-Computer Interfaces: State of the Art and Trends. *IEEE Rev. Biomed. Eng.* **18**, 26–49 (2025).

[R2] Jin, J. et al. MOCNN: A Multiscale Deep Convolutional Neural Network for ERP-Based Brain-Computer Interfaces. *IEEE Trans. Cybern.* **54**, 5565–5576 (2024).

[R3] Chen, X., Wang, Y., Gao, S., Jung, T.-P. & Gao, X. Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface. *J. Neural Eng.* **12**, 046008 (2015).

[R4] Lawhern, V. J. et al. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *J. Neural Eng.* **15**, 056013 (2018).

[R5] Waytowich, N. et al. Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials. *J. Neural Eng.* **15**, 066031 (2018).

[R6] Zhang, J., Gao, S., Zhou, K., Cheng, Y. & Mao, S. An online hybrid BCI combining SSVEP and

EOG-based eye movements. Front. *Hum. Neurosci.* **17**, 1103935 (2023).

Action Taken: Following your comment, we have made some revisions in our manuscript. To ensure clarity, we added more detailed explanations in the revised manuscript, providing a clearer understanding of the entire transmission process. We also added a dedicated discussion on the trade-offs between bit rate, reliability, and security in the revised manuscript.

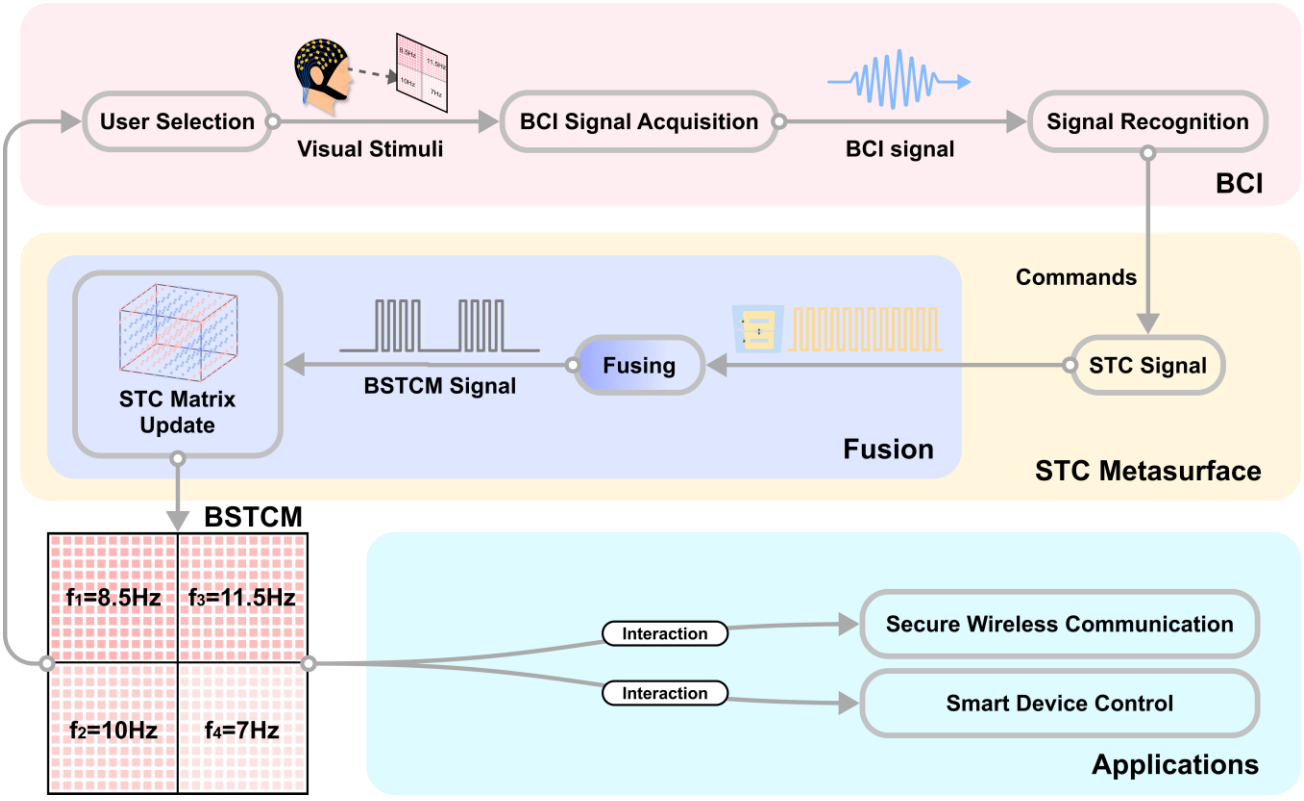

Figure R1. The detailed flow diagram of the BSTCM system. The user generates the BCI signals by gazing at visual stimuli, which are processed and classified into different commands by the machine learning method. These commands are sent to the STC metasurface, which fuses the visual stimuli signals and STC signals to form different BSTCM signals for the secure wireless communications and remote controls of smart devices through the frequency- and space-multiplexed physical channels.

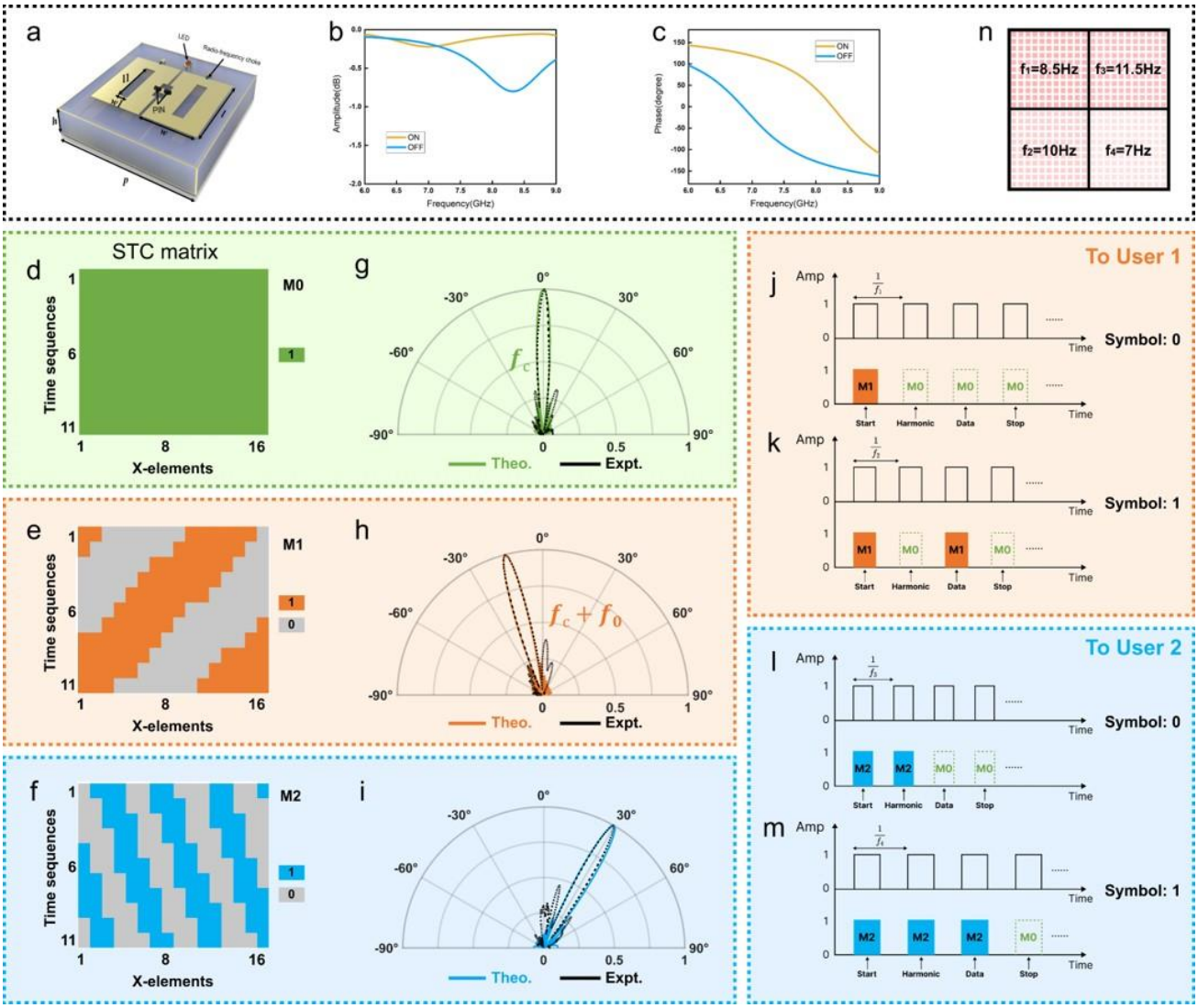

Figure R2. The details of the STC metasurface. a. The geometrical structure of the STC metasurface element. b, c. The reflective amplitude and phase of the element, respectively. d-f. The optimized STC matrices for different scattering angles at different harmonic frequencies. g-i. The far-field results corresponding to the optimized STC matrices. j-m. The encoding schemes for the transmitting symbol "0" or "1" for two users: the symbols "0" and "1" for Users 1 (j and k); the symbols "0" and "1" for Users 2 (l and m). n. The detailed configuration of four visual stimulation regions on the STC metasurface.

[Design of STC metasurface]

…Based on the encoding process, the digital symbol "0" or "1" for User 1 is encoded as the repeated data frames "M1-M0-M0-M0-M1-M0-M0-M0…" or "M1-M0-M1-M0-M1-M0-M1-M0…". Similarly, the digital symbol "0" or "1" for User 2 is encoded as repeated data frames "M2-M2-M0-M0-M2-M2-M0-M0…" or "M2-M2-M0-M0-M2-M2-M0-M0…", respectively.….

[Secure encrypted wireless communication system]

…When the BCI operator (Alice) sequentially transmits the two bitstream sequences of two VSKs to two legal users (Bob and Carol) by the human mind, she freely gazes at the visual stimuli of the four STC metasurface regions based on the properties of the SSVEP-based BCI system. As a result, different EEG signals are generated and classified into four commands, which are then sent to the FPGA to produce the corresponding four BSTCM signals, as illustrated in Fig. 4j–m. Each gaze from the BCI operator corresponds to transmitting a single pixel of the VSK. Due to the neural response time constraints and the required high classification accuracy, the system's data transmission rate is about 0.25 bps, which can be further improved with advanced technology (refer to Supplementary Note 17). By employing the proposed BSTCM system to construct two harmonic frequency-dependent physical channels ($+1^{st}$ and -1^{st} harmonics), two VSKs (VSK1 and VSK2) are sequentially delivered to the respective legal users through these channels. During decryption, the two VSKs corresponding to each secret are XOR-based to recover the original information. The details of encryption and decryption methods can be found in Supplementary Note 12. The target image in question can be decrypted through the XOR operation of two VSKs, realizing a robust level of information security. In the future, the encryption scheme can be expanded to include more users, further enhancing the BSTCM system's security….

[Supplementary Note 17: The Advanced Technologies for Improving Data Transmission Rates]

1. More Precise and Sensitive BCI Electrodes

The evolution of EEG-based BCI systems hinges critically on advancements in electrode technology to achieve higher precision and sensitivity[4]. Current research emphasizes the development of dry electrodes and wireless systems that minimize motion artifacts while maintaining the signal fidelity. The modern wireless EEG devices now support high-density configurations (up to 128 channels) with the sampling rates exceeding 2kHz, enabling ambulatory applications without the constraints of traditional wet electrodes. Emerging innovations are focused on optimizing the electrode-scalp impedance through nanomaterials and flexible substrates, which enhance spatial resolution while ensuring user comfort during prolonged use. For instance, the integration of graphene-based dry electrodes can improve the signal-to-noise ratios (SNR) in motor imagery paradigms by reducing skin-electrode contact resistance. Furthermore, adaptive electrode arrays with embedded signal conditioning circuits are being explored to dynamically compensate for environmental noise and physiological artifacts, particularly in real-world scenarios involving head movement or muscle activity. The miniaturization of low-power, high-gain amplifiers directly into electrode units promises to revolutionize mobile BCIs by enabling robust neural decoding during walking or upper-limb movements. Future directions may incorporate biocompatible hydrogel interfaces and optically transparent electrodes to facilitate simultaneous EEG recordings, thereby capturing complementary hemodynamic and electrophysiological biomarkers for hybrid BCI systems.

2. Smarter and More Efficient Signal Processing Algorithms

Advancement of EEG-based SSVEP-BCI technology is poised for significant acceleration

through the development of smarter and more efficient signal processing algorithms, as evidenced by recent methodological innovations. Deep learning approaches have shown remarkable potentials in addressing the inherent challenges of SSVEP classification, particularly through architectures that optimize feature extraction across multiple spatiotemporal resolutions. The multiscale octave convolution neural network concept, initially proposed for ERP detection[5], presents a promising framework for SSVEP applications through its frequency-component separation strategy, where high-, medium-, and low-frequency features are processed in parallel branches with optimized filter allocation. This architecture's inherent capacity for interactive learning between frequency components could be adapted to enhance SSVEP harmonic analysis while reducing computational overhead. Current research in SSVEP deep learning models reveals a trend toward hybrid architectures that combine temporal and spectral processing, with 72% of surveyed models incorporating multi-branch structures to handle fundamental frequencies and harmonics simultaneously[6]. The integration of filter bank techniques with deep neural networks, as demonstrated in recent studies achieving 94.29% accuracy through optimized frequency-band decomposition[7], suggests further improvements in signal-to-noise ratio management. Moreover, the emergence of lightweight convolutional architectures like EEGNet[8] and Compact-CNN[9], which maintain high performance with reduced parameter counts6, indicates a pathway toward real-time processing optimization. Hybrid paradigm integration, such as the SSVEP-EOG combination achieving 117.05 bits/min information transfer rate through time-space frequency conversion[10], exemplifies how multimodal signal fusion could enhance system robustness. Future algorithmic developments may leverage emerging techniques in transfer learning4 and manifold classification to address inter-subject variability, potentially building upon open-source toolbox innovations that facilitate rapid algorithm prototyping. The systematic optimization of hyperparameters in deep learning models, as evidenced by recent studies achieving 92.15% accuracy through Adam optimizer-driven training with spatial dropout[11], underscores the importance of computational efficiency in algorithm design. These collective advancements in multi-scale feature extraction, hybrid architectures, and computational optimization strategies position SSVEP-BCI systems for substantial performance improvements in both speed and reliability.

3. Higher Metasurface Control Rate

The advancement of metasurface control rates critically depends on further breakthroughs in transistor performance, as evidenced by recent developments in high-speed and high-density integrated circuits. Future research could focus on optimizing material systems and device architectures to overcome existing limitations. By addressing challenges in contact resistance reduction, dielectric optimization, and process compatibility, next-generation transistors could achieve switching frequencies exceeding current state-of-the-art benchmarks. These advancements will not only enhance metasurface control rates but also enable unprecedented integration densities and energy efficiencies for emerging applications in adaptive optics, biomedical sensing, and ultrafast communication systems. The control rate of the metasurface can be significantly improved by employing transistors with higher switching speeds, thereby reducing the response time of individual unit cells. Furthermore, the integration of high-speed FPGA architectures significantly enhances the metasurface's control rate, thereby enabling ultra-fast reconfiguration of electromagnetic properties.

Comment 1.2: *It appears that all information transfer is initiated by the LED lights. Does the host user have the ability to freely transmit data of their choice? If not, what is the justification for using a Brain-Machine Interface (BMI)? In this case, wouldn't a standard ST metasurface multi-user scheme be sufficient?*

Response: Thank you very much for your insightful comments. We would like to clarify the following points regarding to user control, the role of LED lights, and the justification for employing a Brain-Machine Interface (BMI) in our system.

1. User-Initiated Data Transmission

The host user has the ability to freely transmit data of their choice using the BSTCM system. During secure communication, the host user sequentially sends two VSKs to two users by the proposed BSTCM system. In the smart device control experiment, the host user can freely choose the smart device to be controlled by staring at the STC metasurface. We would like to emphasize that while the LED lights play a role in initiating the information transfer by inducing SSVEP signals of the human brain, the control of the information transmission is indeed dependent on the brain activity of the host user. By gazing at flashing stimuli of different frequencies associated with different regions of the STC metasurface, the host user's cognitive intent leads to the generation of specific EEG signals. These EEG signals are subsequently processed by the BCI system to classify the corresponding commands. The commands are used to fuse the visual stimuli signals and STC signals with the help of an FPGA, resulting in the formation of four BSTCM signals, as depicted in Figure R1(j-m). These signals control the STC metasurface, enabling the transmission of different messages to other users. The primary motivation for integrating the BMI in our system is to allow direct mind-to-device communication, bypassing the need for traditional input devices such as keyboards or mice.

2. Comparison with a Standard STC Metasurface Multi-User Scheme

The combination of BCI with STC metasurfaces offers a unique, non-invasive means of communication that could be highly beneficial in applications where traditional user interfaces are impractical or undesirable (e.g., for individuals with disabilities or in hands-free operation environments). In addition, STC metasurfaces alone could facilitate multi-user communication, the integration of BCI enhances flexibility, interactivity, and personalized communication. The BCI

provides a direct interface with the human brain, enabling users to control the communication process through their cognitive intent, while the metasurface ensures secure transmission by utilizing frequency-dependent physical channels and encryption techniques. This combination of human-computer interaction and advanced security mechanisms offers a level of control and protection that surpasses the capabilities of a standard STC metasurface scheme. We hope this clarification addresses the reviewer's concerns.

Action Taken: Following your comment, we have made some revisions in our manuscript.

[Design of STC metasurface]

> …Hence, the BCI operator can freely transmit different information by staring at different visual stimuli of the STC metasurface according to the above-mentioned encoding method.

[Mind-controlled smart devices]

> …Initially, distinct EEG signals can be generated when the BCI operator stares at different visual stimuli of the STC metasurface. The EEG signals are acquired and subsequently processed by the BSTCM system for command recognition…

Comment 1.3: *Furthermore, a BER of 30% is quite high for the legal user (Fig.R2), which raises concerns about the system's reliability and practical usability.*

Response: Thank you very much for your insightful comments. We sincerely apologize for any confusion in presenting the figure content. To clarify, as shown in Figure R3c, the BER of 30% refers to the bit error rate of the data intercepted by the illegal user during the wireless transmission of VSK2 when the illegal user is aware of the harmonic frequency and is positioned close to legal user 2. Importantly, when the illegal user is located at other spatial positions far from legal user 2, the BER approaches 50%, demonstrating that it becomes increasingly difficult to eavesdrop on correct information of VSK2. Additionally, as illustrated in Figure R3d, a BER of approximately 30% occurs when the illegal user, knowing the harmonic frequency, is positioned in the 35° spatial direction near legal user 2. However, when the illegal user attempts to eavesdrop on the information of VSK2 at different harmonic frequencies, the BER approaches 50%, further reinforcing the system's reliability against eavesdropping. More critically, the successful retrieval of transmitted information necessitates the simultaneous acquisition of both VSK1 and VSK2, substantially increasing the complexity and

difficulty for illegal users to accurately reconstruct the original message and proving the system's reliability and practical usability. This dual-channel requirement enhances security and ensures that an eavesdropper cannot easily decipher the communication.

Action Taken: Following your comment, we have added the detailed description in the main text and supplementary materials of the revised manuscript.

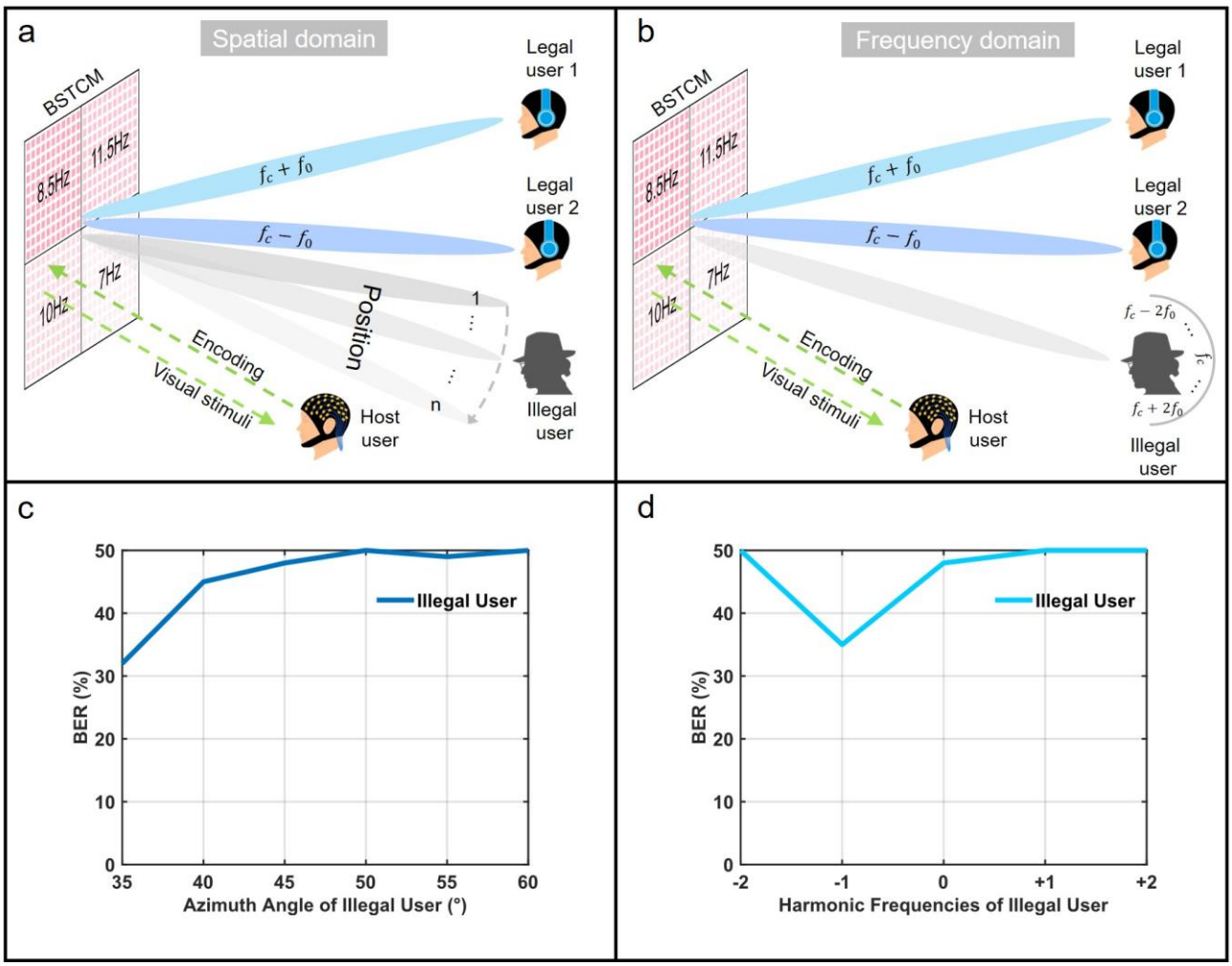

Figure R3. The security measurement of the proposed secure wireless communication scheme. a. The diagram of security test scenarios for illegal users in different spatial locations from 35° to 60°. b. The diagram of security test scenarios for illegal users in different harmonic frequencies from -2nd to +2nd order. c. The measured results for illegal users eavesdropping at -1st harmonic frequency in different spatial positions from 35° to 60°. d. The measured results for illegal users eavesdropping at a spatial position of 35° in different harmonic frequencies from -2nd to +2nd order.

[Supplementary Note 15: The Security Performance of the Proposed Secure Wireless Communication]

…Firstly, we measured the BER results of the illegal user with -1st harmonic frequency in different spatial locations from 35° to 60° in the BSTCM-based wireless communication system. The BERs are nearly 50% shown in Supplementary Figure 26c. The illegal users with different harmonic frequencies from -2nd to +2nd order have been tested when they are located at 35° spatial direction. Supplementary Figure 26c demonstrated that BERs are nearly 50%....

[Secure encrypted wireless communication system]

…Furthermore, the BERs of the illegal user are measured to be approximately 50% when the illegal user eavesdrops separately from both the spatial and frequency domains (refer to Supplementary Note 15 for additional details).….

REVIEWER #3

We thank this reviewer for the positive feedback regarding to our manuscript. In particular, the reviewer stated that "*The writing of the resubmitted article is much easier to understand…*", and gave several insightful and valuable comments, which are very helpful to further improve the quality of our manuscript. We have addressed all comments carefully, and all changes are listed as follows.

Comment 3.1: *The writing of the resubmitted article is much easier to understand, at least up to the section on secure communication. However, this part remains highly confusing, and the provided corrections and accompanying letter do not sufficiently clarify the content. Once again, I spent a significant amount of time—nearly an entire day—trying to understand the coding process described.*

Response: Thank you very much for your professional comments. To address your concerns, we have significantly revised the section on secure communication to enhance clarity, structure, and readability. The entire communication process is described as follows:

EEG signals recognition: The detailed flow diagram of the BSTCM system is illustrated in Figure R4. The system begins with flickering visual stimuli, designed to elicit specified SSVEP-based EEG signals in the operator's brain. These original EEG signals are then captured by BCI amplifier devices and displayed in Figure R5a. The raw EEG signals are then preprocessed to feature graphs, as shown in Figure R5b. A convolutional neural network (CNN) of Figure R5c is utilized to classify the feature maps derived from EEG signals. This approach is widely adopted in EEG-based BCI systems owing to the inherent complexity and non-stationary nature of EEG signals. As shown in Figure R5d, the classifier outputs confidence scores for each stimulus frequency, with the highest score being mapped to control commands, hence realizing the interaction. This end-to-end pathway translates neural activity into functional outputs, enabling non-invasive human-computer interaction. Once the interaction commands are identified, they are transmitted to an FPGA. The FPGA serves as the central control module, integrating the visual stimulation signals with the corresponding STC matrices to generate a BSTCM signal. This signal drives the metasurface, enabling it to perform dual functions: providing the visual stimuli required for interaction and supporting advanced applications such as encrypted wireless communication systems or remote smart device control. This innovative approach enhances system efficiency and compactness by unifying visual stimulation, signal processing, and application control within a single platform.

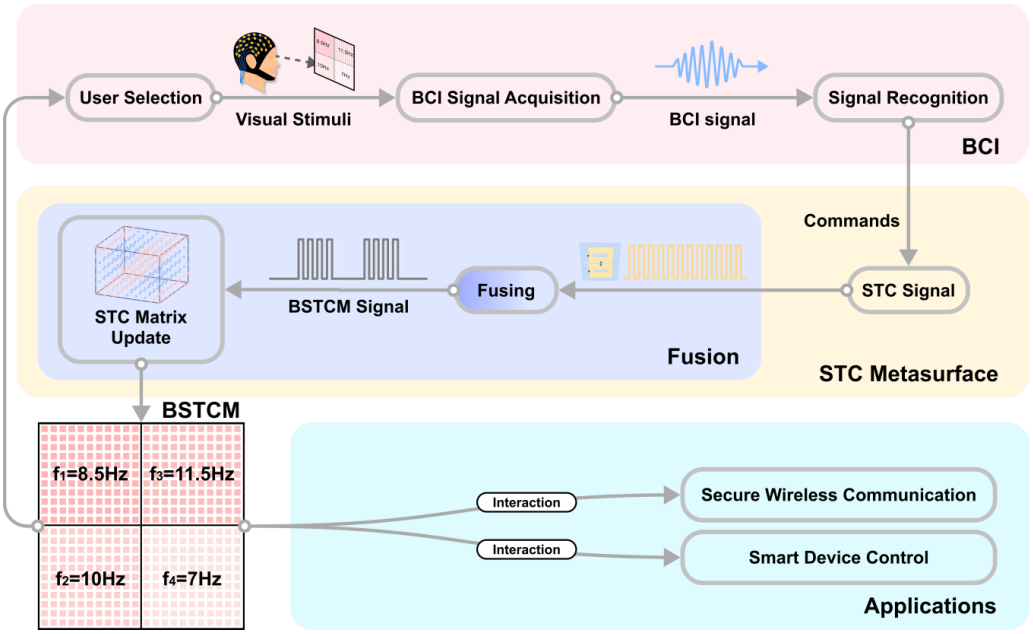

Figure R4. The detailed flow diagram of the BSTCM system. The user generates the BCI signals by gazing at visual stimuli, which are processed and classified into different commands by the machine learning method. These commands are sent to the STC metasurface, which fuses the visual stimuli signals and STC signals to form different BSTCM signals for the secure wireless communications and remote controls of smart devices through the frequency- and space-multiplexed physical channels.

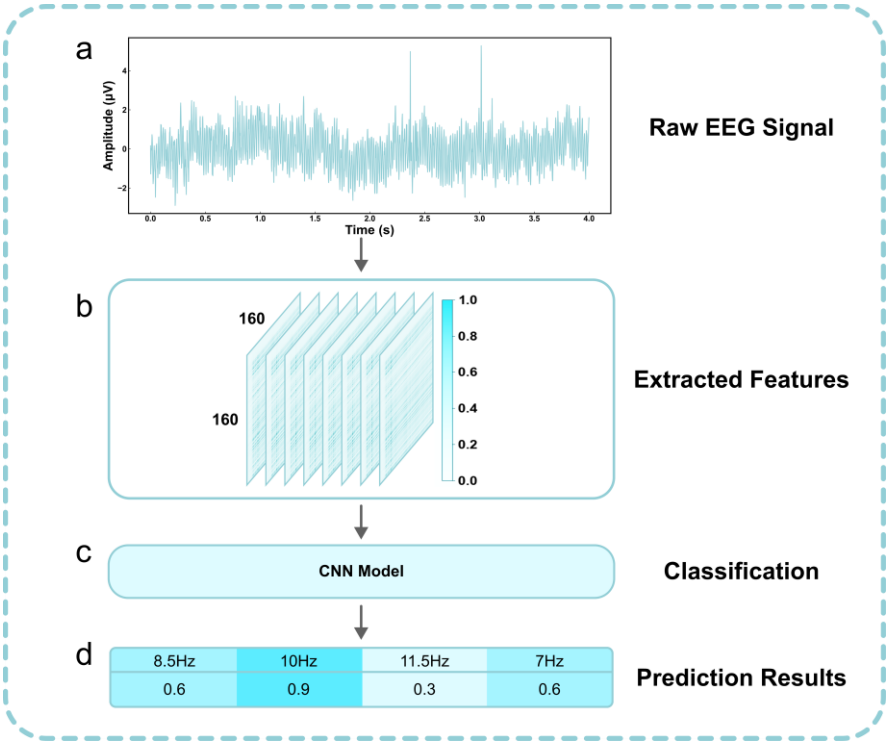

Figure R5. The detailed schematic diagram of EEG signals recognition. a. The raw EEG signals are generated by gazing at visual stimuli. b. The raw EEG signals are preprocessed as feature graphs. c. The preprocessed feature graphs are then input into a CNN model for classification. d. The prediction results of the raw EEG signals from the CNN model.

Cryptography encryption method: To enhance clarity and simplify the system, we have removed certain non-essential processes, making the secure communication framework more intuitive. The transmitted information is first encrypted using the method illustrated in Figure R5. As depicted in Figure R6, the secret image "H" is represented as a 5×5 grayscale matrix, where black pixels correspond to digital “1” and white pixels correspond to digital “0”. The encryption process is based on an XOR-based coding scheme. Initially, a randomly generated VSK1 matrix (5×5 pixels) is created. The corresponding VSK2 matrix, with the same dimensions, is then computed using the XOR operation between VSK1 and the original secret image “H”. Consequently, each pixel of “H” is encrypted into two shared sub-pixels, which are subsequently combined to form two VSKs. Crucially, neither VSK1 nor VSK2 alone contains sufficient information to reconstruct the original secret information, ensuring the security of the transmission. Furthermore, Figure R7 demonstrates the encryption process for additional secret images. Using the same VSK1, we calculate three distinct VSK2 matrices, as shown in Figure R7a–c. Additionally, we generate three different random VSK1 matrices and calculate the corresponding VSK2 matrices, as illustrated in Figure R7d–f. These results further validate the robustness and flexibility of the encryption scheme.

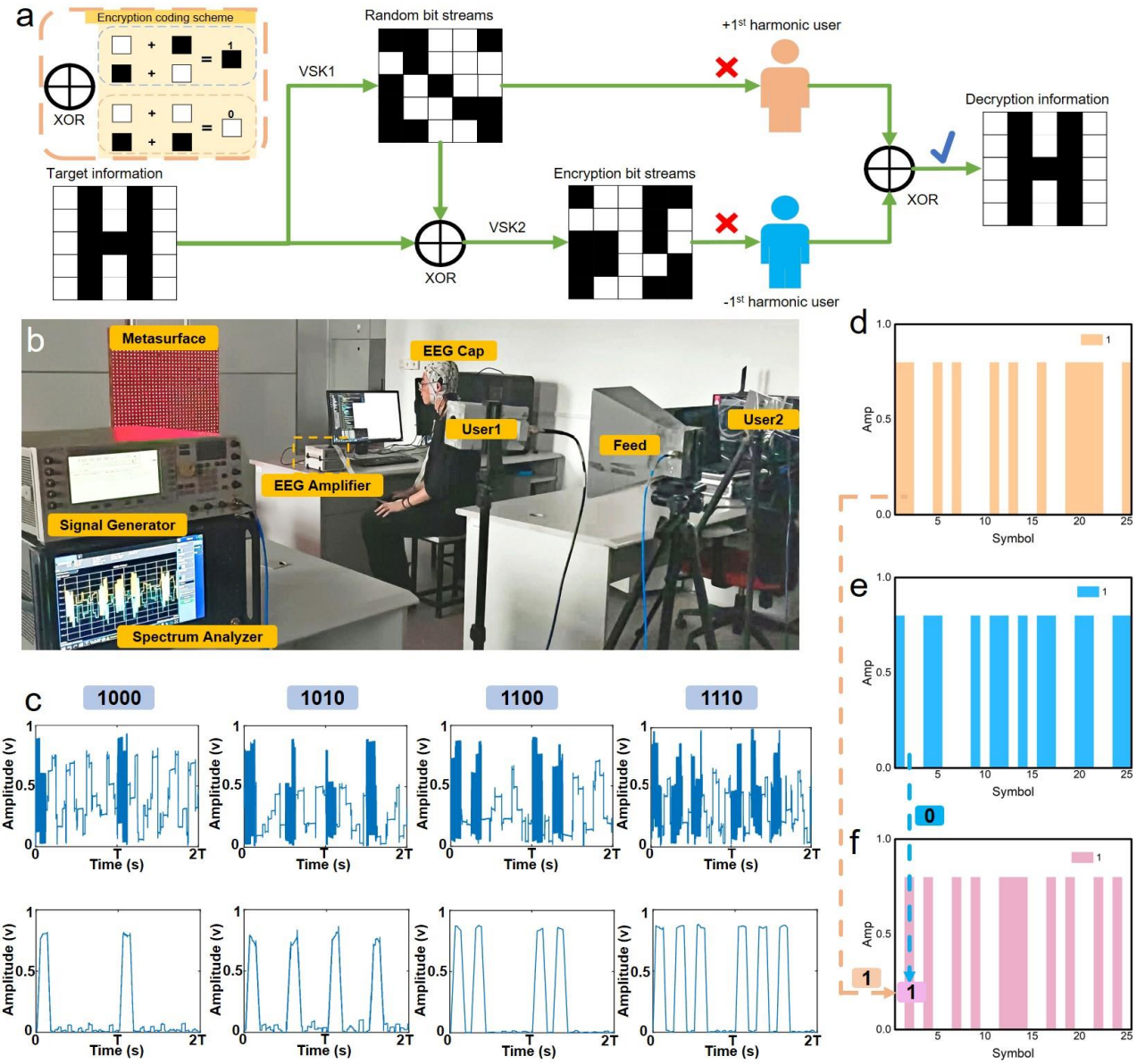

Figure R6. The encrypted wireless communication system based on the BSTCM platform. a. The encrypted encoding scheme for the secret image "H". b. The experimental scenarios of the encrypted wireless communication system, in which the operator equipped with EEG cap transmits the encrypted information to two users via the BSTCM platform, respectively. c. The receiving signals corresponding to the encoding scheme in Fig 3(j-m). d, e. The decoded information of VSK1 and VSK2 from the receiving signals based on the encrypted coding scheme. f. The correct transmitting information is decoded by the extracted VSK1 and VSK2.

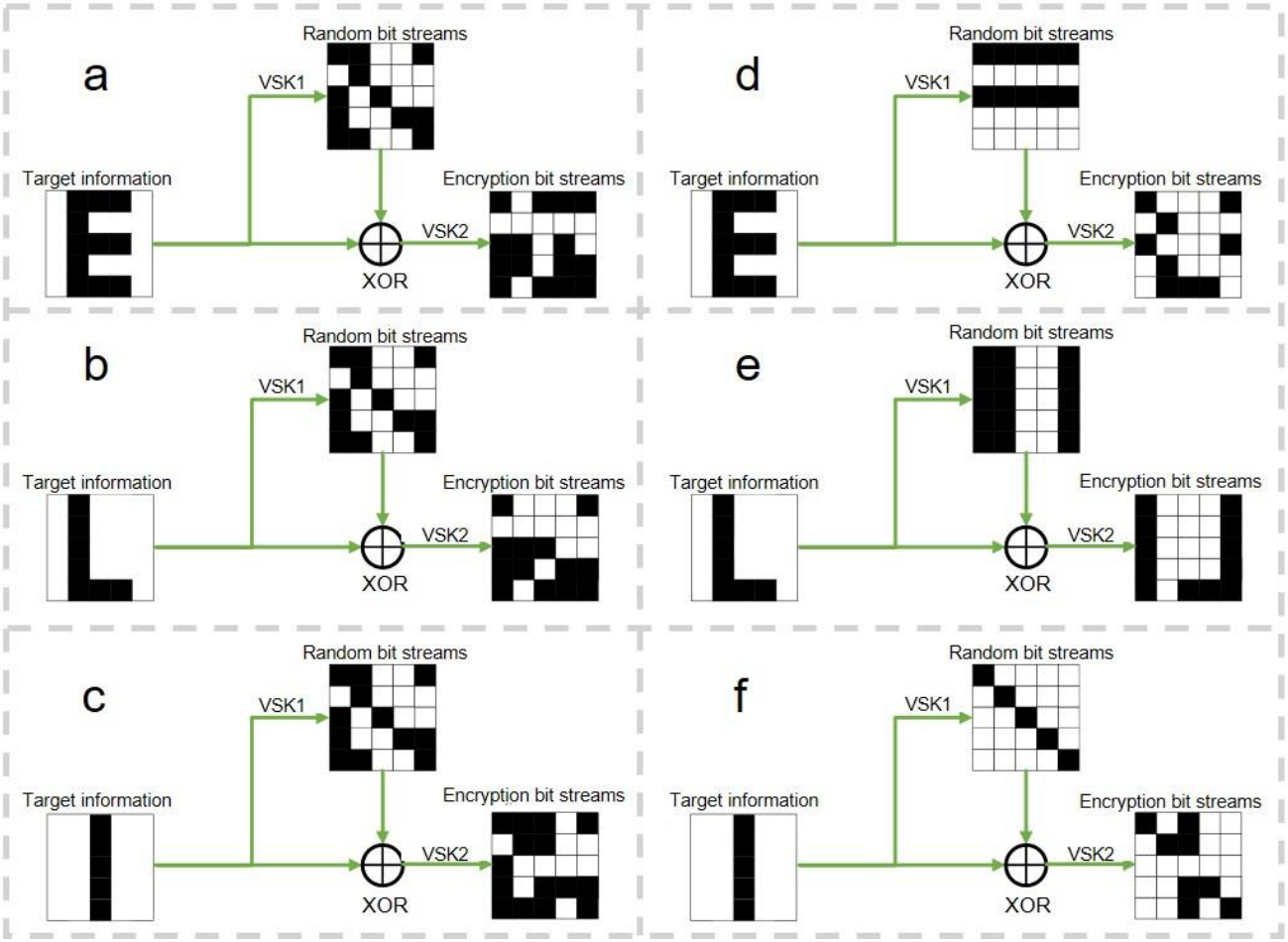

Figure R7. The encryption coding process of three secret letter images "E", "L", and "I". a-c. According to the encryption encoding scheme, the same VSK1 is utilized to generate the corresponding VSK2 for three secret letter images "E", "L", and "I". d-f. The corresponding VSK2 of three secret letter images can be calculated using randomly generated different VSK1.

Encoding process: Based on the proposed BSTCM system, the host user can gaze at different partitions of the STC metasurface to transmit two distinct ciphertexts (VSK1 and VSK2) to two separate users, and the encoding schemes are illustrated in Figure R8j-m. As shown in Figure R8n, we demonstrate the detailed configuration of four visual stimulation regions on the STC metasurface. Specifically, the left partition of the STC metasurface is designated for transmitting symbols "0" and "1" of VSK1 to User 1, while the right partition is responsible for transmitting symbols "0" and "1" of VSK2 to User 2. The modulation scheme of the proposed wireless communication system is Amplitude Shift Keying (ASK) modulation. The STC matrix 1 (M1) of Figure R8e is input into the STC metasurface through an FPGA, which controls the electromagnetic waves to construct a $+1^{st}$ harmonic communication channel with a beam direction of -15°. The STC matrix 2 (M2) is utilized to construct

a -1st harmonic communication channel, with a beam direction of +30°. The STC matrix 0 (M0) is utilized to represent bit "0", with a beam direction of 0°. Therefore, two users can not receive any energy when the STC matrix switches to M0. In order to integrate visual stimuli and electromagnetic regulation, different STC matrix signals are loaded into the visual stimuli signals to support both visual stimuli and information transmission. Hence, the symbol "0" or "1" of VSK1 is represented by periodically repeated data frames: "M1-M0-M0-M0-M1-M0-M0-M0…" or "M1-M0-M1-M0-M1-M0-M1-M0…". The fused BSTCM signal cycles periodically, with each data frame consisting of four bits, where the third bit of each frame serves as the transmitted data bit and the remaining bits are used as check bits. Similarly, the symbol "0" or "1" of VSK2 is represented by periodically repeated data frames: "M2-M2-M0-M0-M2-M2-M0-M0…" or "M2-M2-M2-M0-M2-M2-M2-M0…".

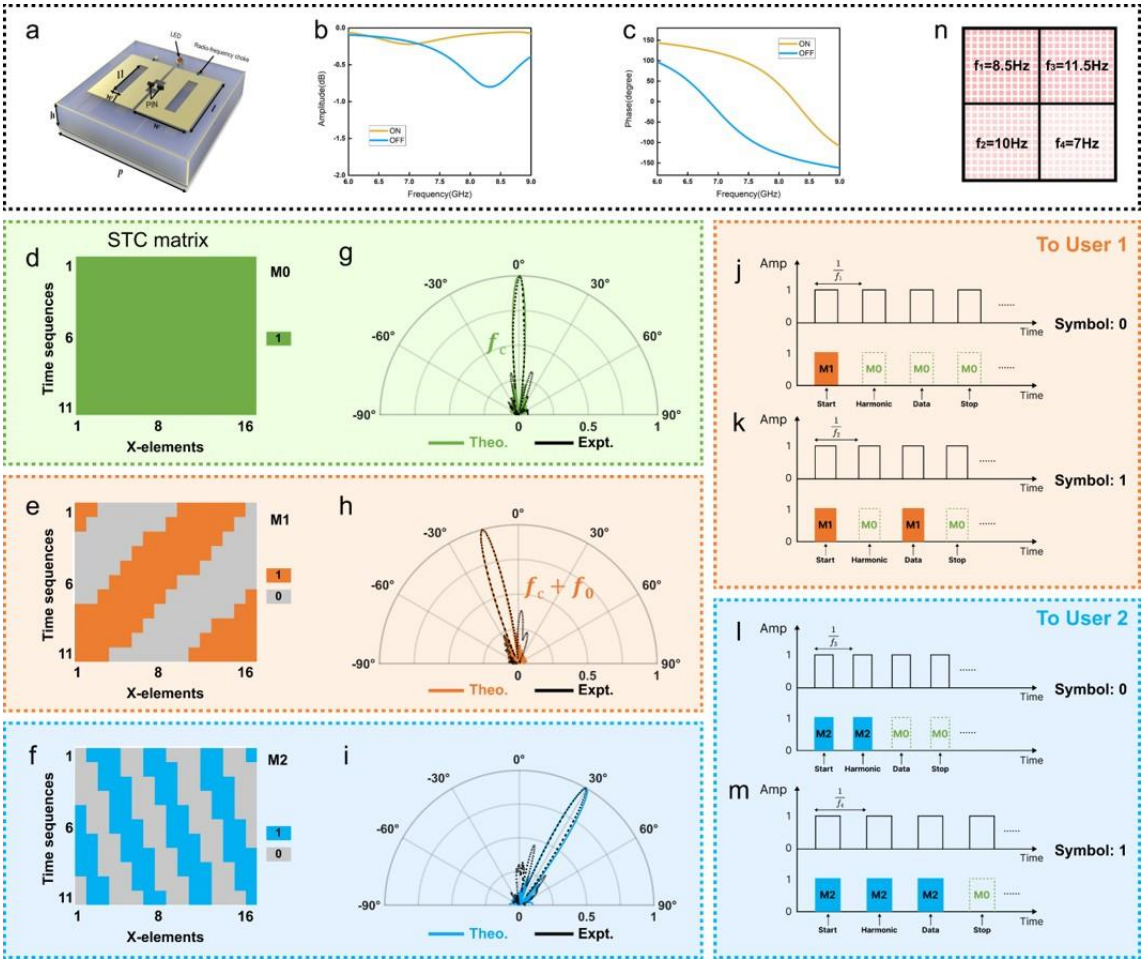

Figure R8. The details of the STC metasurface. a. The geometrical structure of the STC metasurface element. b, c. The reflective amplitude and phase of the element, respectively. d-f. The optimized STC matrices for different scattering angles at different harmonic frequencies. g-i. The far-field results corresponding to the optimized STC matrices. j-m. The encoding schemes for the transmitting symbol "0" or "1" for two users: the symbols "0" and "1" for Users 1 (j and k); the symbols "0" and "1" for Users 2 (l and m). n. The detailed configuration of four visual stimulation regions on the STC metasurface.

Wireless communication process: The complete transmitting and receiving process is shown in Figure R9-11. Firstly, the host user generates corresponding EEG signals by gazing at different STC metasurface partitions according to the content of two VSKs, these EEG signals are subsequently processed and classified into the corresponding control commands. The commands are then input into the FPGA, which integrates the STC signals and visual stimuli signals according to the encoding scheme illustrated in Figure R8j-m. The constructed harmonic communication channels then securely transmit the encrypted information to the designated users represented by two standard horn antennas. A commercial spectrum analyzer is used to receive information sent by the host user. The received information is demodulated and transformed by inverse fast Fourier transform, as shown in Figure R6c. As demonstrated in Figure R6d and R6e, the data bits are sequentially extracted from the received information and combined into two VSKs (VSK1 and VSK2), respectively. According to the encryption coding scheme, the secret information can be recovered by XOR operation of two VSKs, as shown in Figure R11.

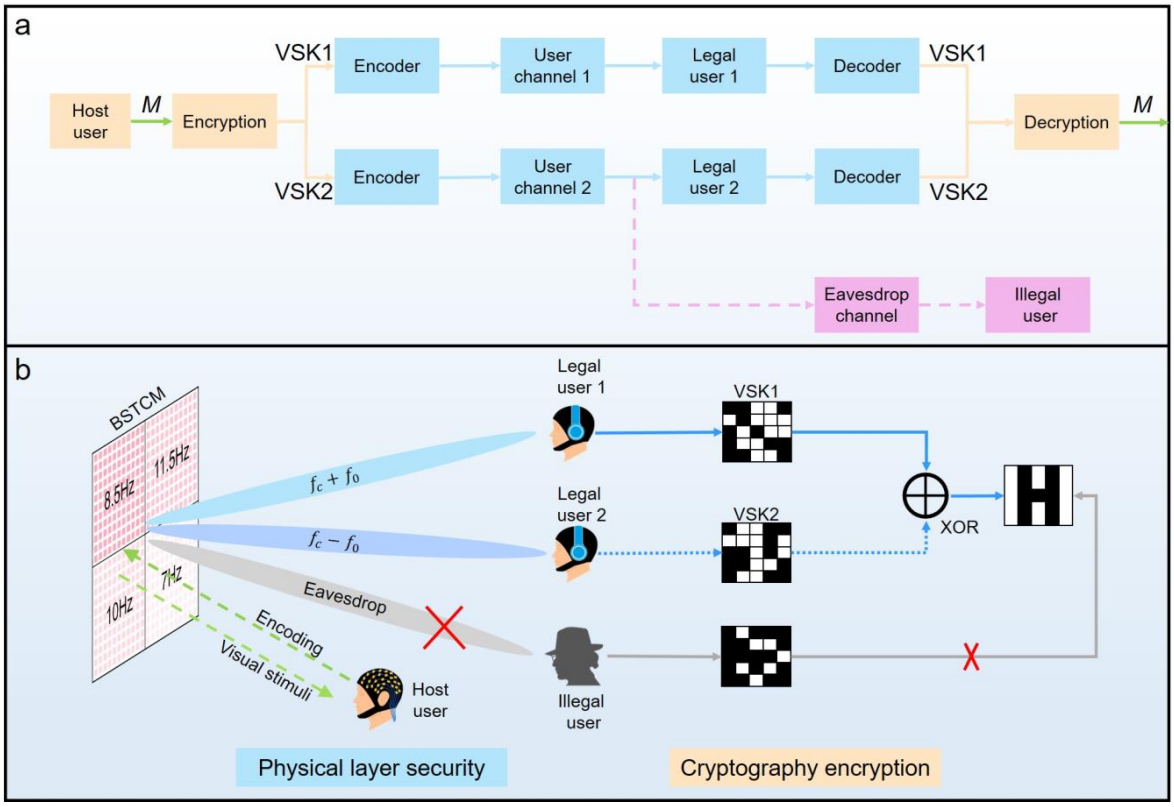

Figure R9. The proposed secure wireless communication scheme. a. The detailed flowchart of the proposed security model. b. The schematical diagram of the presented BSTCM-based secure wireless communication system.

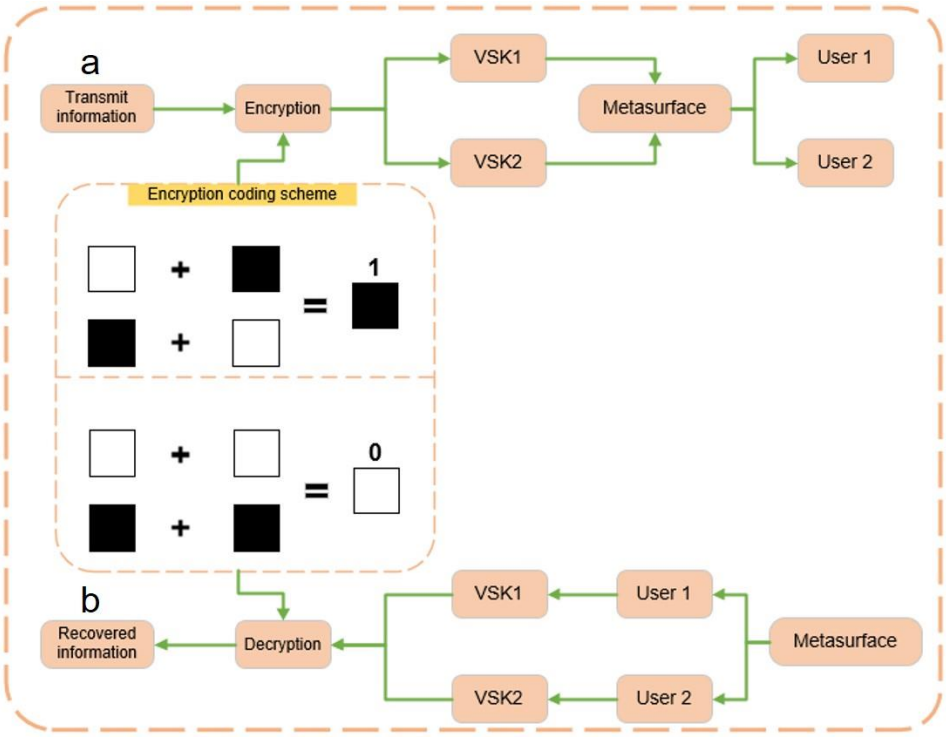

Figure R10. The presented wireless communication process. a. The target information is encrypted to send to two legal users. b. The target information is decrypted by the received information from two legal users.

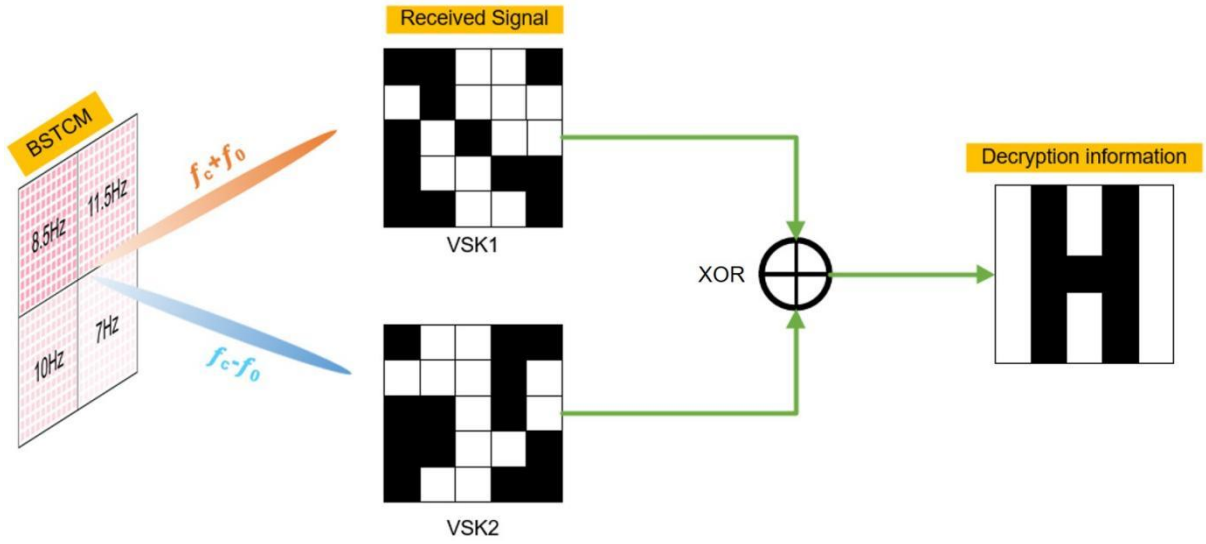

Figure R11. The decrypted coding process of the secret image "H" from two legal users at $+1^{st}$ and -1^{st} harmonic frequency according to two shared VSKs.

Action Taken: Following your comment, we have added the detailed description in the main text and supplementary materials of the revised manuscript.

[Design of STC metasurface]

…Based on the encoding process, the digital symbol “0” or “1” for User 1 is encoded as the repeated data frames “M1-M0-M0-M0-M1-M0-M0-M0…” or “M1-M0-M1-M0-M1-M0-M1-M0…”. Similarly, the digital symbol “0” or “1” for User 2 is encoded as repeated data frames “M2-M2-M0-M0-M2-M2-M0-M0…” or “M2-M2-M0-M0-M2-M2-M0-M0…”, respectively.….

[Secure encrypted wireless communication system]

…In the proposed secure encrypted wireless communication system, the transmitted information is firstly encrypted into two ciphertexts. These ciphertexts are then securely sent to two designated users via the BSTCM system, which are seamlessly controlled by human intention. We present an encryption strategy designed to facilitate the secure wireless communication influenced directly by the human intention. Under this approach, the original information is recovered by XOR-based stacking of two visual shared keys (VSKs), wherein each pixel of the secret message is encrypted into two sub-pixels for two distinct users. As illustrated in Fig. 5a, VSK1 is randomly generated and VSK2 is subsequently derived through the XOR operation. Notably, neither share alone contains any discernible details regarding to the original secret, thereby ensuring that the disclosure of a single VSK does not compromise the confidentiality of the underlying message. As an illustrative example in Fig. 5a, the secret image “H”, consisting of a grid of 5×5 black pixels (digital “1”) and white pixels (digital “0”), is encrypted into two VSKs composed of 5×5 sub-pixels based on the presented encryption method, where VSK1 is randomly generated and VSK2 is subsequently derived through the XOR operation. Consequently, neither VSK1 nor VSK2 alone contains sufficient information to reveal the original secret information. When the BCI operator (Alice) sequentially transmits the two bitstream sequences of two VSKs to two legal users (Bob and Carol) by the human mind, she freely gazes at the visual stimuli of the four STC metasurface regions based on the properties of the SSVEP-based BCI system. As a result, different EEG signals are generated and classified into four commands, which are then sent to the FPGA to produce the corresponding four BSTCM signals, as illustrated in Fig. 4j–m. Each gaze from the BCI operator corresponds to transmitting a single pixel of the VSK. Due to the neural response time constraints and the required high classification accuracy, the system’s data transmission rate is about 0.25 bps, which can be further improved with advanced technology (refer to Supplementary Note 17). By employing the proposed BSTCM system to construct two harmonic frequency-dependent physical channels (+1st and -1st harmonics), two VSKs (VSK1 and VSK2) are sequentially delivered to the respective legal users through these channels. During decryption, the two VSKs corresponding to each secret are XOR-based to recover the original information. The details of encryption and decryption methods can be found in Supplementary Note 12. The target image in question can be decrypted through the XOR operation of two VSKs, realizing a robust level of information security. In the future, the encryption scheme can be expanded to include more users, further enhancing the BSTCM system’s security.

[Supplementary Note 12: The Harmonic-Based Encrypted Wireless Communication Scheme]

…Firstly, the encryption coding scheme is designed based on the XOR operation. The digital "1" of the secret image can be decrypted when an exclusive OR (XOR) relationship is set between two information blocks. Whereas, the digital "0" of the secret image can be decrypted when an exclusive NOR (XNOR) relationship is set between two information blocks. The secret image can be encrypted as two information block matrices (VSK1 and VSK2). According to the encryption method, two VSKs will be sent to two different users with different harmonic frequencies, respectively. As a result, every single secret is divided into two matrices so that the effective information can be hidden as two VSK1 and VSK2. In this way, the contents of transmitting single VSKs are completely random and meaningless, which makes either VSK1 or VSK2 strongly robust against different eavesdroppers. Then, two random VSK1/VSK2 are sent to two users with two harmonic frequencies by the constructed secure encrypted wireless communication system based on the BSTCM platform….

[Supplementary Note 16: Schematic Diagram of EEG Signals Recognition]

As shown in Supplementary Figure 27, the SSVEP-BCI in our BSTCM system operates through a sequential signal-processing pipeline. Initially, visual stimuli flickering at distinct non-harmonic frequencies (7 Hz, 8.5 Hz, 10 Hz, 11.5 Hz) evoke frequency-locked steady-state visual evoked potentials (SSVEPs) in the user's occipital cortex. These raw neural response signals are captured via an EEG amplifier. The raw signals are then transformed into feature graphs that encode correlations between incoming SSVEP spectral profiles and precomputed reference templates derived from baseline data. A convolutional neural network (CNN) is employed to classify the feature graphs, a common practice in EEG-based BCI systems due to the complexity of EEG signals. The classifier outputs confidence scores for each stimulus frequency, with the highest score being mapped to actionable commands (e.g., device control or data transmission) through downstream hardware integration. This end-to-end pathway translates neural activity into functional outputs, enabling non-invasive human-computer interaction.

Comment 3.2: *Visual Secret Sharing (VSS): Why is this termed "visual"? Beyond the secure transmission section, the role of visual stimuli is not explained or justified.*

Response: Thank you very much for your professional comments. In our work, the concept of visual stimuli and visual secret sharing (VSS) are distinct and not directly related. To prevent any ambiguity, we will remove the term "VSS" from the manuscript. Additionally, we will clarify the role of visual stimuli in the context of our study and ensure its justification is clearly explained.

1. Visual Secret Sharing

The term "Visual Secret Sharing (VSS)" is used because the decryption process is based on human visual perception rather than computational decryption[R1-R3]. Unlike traditional cryptographic methods, which require mathematical algorithms for recovery, VSS relies on overlapping shares to reveal the

secret information visually. This feature ensures secure and intuitive decryption, making it highly relevant for applications where human verification is essential. To simplify the system and make the secure transmission process clearer, we carefully revised the relevant statements. As demonstrated in Figure R6, the secret image "H" is a grayscale matrix composed of 5×5 black pixels (digital "1") and white pixels (digital "0"). The encryption coding scheme of XOR operation is presented to encrypt the secret information. Hence, each pixel of the secret image "H" is encrypted into two shared sub-pixels, which are ultimately combined to form VSK1 and VSK2. Importantly, neither VSK1 nor VSK2 alone contains sufficient information to reveal the original secret information. The original information can only be accurately retrieved by performing an XOR operation between VSK1 and VSK2. The VSS encryption mechanism is specifically designed for secure information transmission. When the host user intends to transmit VSK1 and VSK2 to other users, the signal transmission is achieved by directing their gaze toward the visual stimuli of varying frequencies emitted by the STC metasurface. Consequently, the VSS mechanism is only used to the secure transmission section.

2. Visual Stimuli

SSVEPs arise from synchronized neural oscillations in the visual cortex when repetitive visual stimuli, such as flickering lights or patterns, are presented at specific frequencies, typically within the 5–60 Hz range[R4]. These oscillations are driven by the entrainment of neural populations in the extrastriate visual cortex, reflecting a sensory gain-control mechanism that amplifies neural responses to attended stimuli10. The necessity of visual stimulation lies in its ability to evoke frequency-specific cortical facilitation, which serves as the primary biomarker for decoding user intent in SSVEP-BCIs. Without the visual stimulation, the distinct, sustained oscillatory patterns required for reliable signal detection and classification would not manifest, rendering the system inoperative[R5]. Different SSVEP-based EEG signals are induced to be generated when users stare at visual stimuli of different frequencies on the STC metasurface. Furthermore, EEG signals are recognized to the corresponding control commands, which then control the STC metasurface to achieve information transmission.

In summary, the term "Visual Secret Sharing" is justified by its reliance on human visual perception for decryption, and its application in this work is primarily focused on secure transmission. Furthermore, the mechanism by which visual stimuli elicit specific EEG responses is a central theme

that underpins the entire framework proposed in this study. Thank you for your valuable feedback, which has significantly enhanced the clarity and depth of our manuscript.

References:

[R1] Feng, J. B., Wu, H. C., Tsai, C. S., Chang, Y. F. & Chu, Y. P. Visual secret sharing for multiple secrets. *Pattern Recognition* **41**, 3572-3581 (2008).

[R2] Zhu, H. Y., Hieu, N. Q., Hoang, D. T., Nguyen, D. N. & Lin, C.-T. New visual secret sharing schemes using probabilistic method. *Pattern Recognition Letters* **25**, 481-494 (2004).

[R3] Chen, T. H., et al. Visual secret sharing by random grids revisited. *Pattern Recognition* **42**, 2203-2217 (2009).

[R4] Müller, M. M., Teder-Sälejärvi, W. & Hillyard, S. A. The time course of cortical facilitation during cued shifts of spatial attention. *Nature Neuroscience.* **1**, 631–634 (1998).

[R5] Wang, Z. et al. Conformal in-ear bioelectronics for visual and auditory brain-computer interfaces. *Nature Communication* **14**, 4213 (2023).

Action Taken: Following your comment, we have deleted the corresponding description in the main text and supplementary materials of the revised manuscript. In addition, we have added explanations of visual stimuli beyond the secure transmission section.

[Introduction]

> …Specifically, the transmitted information is encrypted into two ciphertexts, which are respectively sent to two users through two independent harmonic frequency channels constructed by the BSTCM system….

[System configuration]

> …The visual stimulation is crucial for eliciting the distinct oscillatory patterns necessary for user-intent decoding; and without it, the BCI system would be inoperative.….
> …The target flickering stimuli, characterized by four distinct frequencies (8.5Hz, 10Hz, 11.5Hz, and 7Hz), are supported by the STC metasurface integrated with LEDs….
> …In the encrypted wireless communication system, the XOR-based encryption encoding method is employed to encrypt the target information[34-36]….

[Secure encrypted wireless communication system]

> …We present an encryption strategy designed to facilitate secure wireless communication influenced directly by human intention. Under this approach, the original information is recovered by XOR-based stacking of two visual shared keys (VSKs), wherein each pixel of the secret message is encrypted into two sub-pixels for two distinct users.….

[Mind-controlled smart devices]

> …Initially, distinct EEG signals can be generated when the BCI operator stares at different visual stimuli of the STC metasurface. The EEG signals are acquired and subsequently processed by the BSTCM system for command recognition….

[Supplementary Note 12: The Harmonic-Based Encrypted Wireless Communication Scheme]

> …The proposed high-security model of the BSTCM-based wireless communication system is shown in Supplementary Figure 25, which integrates cryptographic encryption method and physical layer security….

Comment 3.3: *Transmission of VSK1 and VSK2: To my understanding, VSK1 and VSK2 are transmitted to user 1 and user 2 using the +1 and -1 harmonics, respectively, as per the coding schemes illustrated in Figures 4j to 4m. However, this is not explicitly stated in the text.*

Response and Action Taken: Thank you very much for your valuable feedback. We appreciate the opportunity to clarify the transmission mechanism of VSK1 and VSK2. As stated in the response to Comment 3.1, VSK1 and VSK2 are transmitted to User 1 and User 2, respectively, via the +1 and -1 harmonic frequencies beams, following the encoding schemes illustrated in Figure R8j-m. To improve clarity, we have explicitly stated this process in the revised manuscript. Specifically, the STC metasurface modulates the spatial and frequency domain characteristics of the transmitted signals, ensuring that VSK1 and VSK2 are independently encoded onto the +1 and -1 harmonic channels, respectively. We sincerely appreciate your detailed review and have taken your feedback into account to enhance the manuscript's clarity.

[Secure encrypted wireless communication system]

…When the BCI operator (Alice) sequentially transmits the two bitstream sequences of two VSKs to two legal users (Bob and Carol) by the human mind, she freely gazes at the visual stimuli of the four STC metasurface regions based on the properties of the SSVEP-based BCI system. As a result, different EEG signals are generated and classified into four commands, which are then sent to the FPGA to produce the corresponding four BSTCM signals, as illustrated in Fig. 4j–m. Each gaze from the BCI operator corresponds to transmitting a single pixel of the VSK….
…By employing the proposed BSTCM system to construct two harmonic frequency-dependent physical channels ($+1^{st}$ and -1^{st} harmonics), two VSKs (VSK1 and VSK2) are sequentially delivered to the respective legal users through these channels….

Comment 3.4: *Harmonics: The frequency f0=1kHz is much smaller than the coherence frequency of classical propagation environments. Therefore when the STC metasurface beams in a specific direction with a commuting frequency fc, the signal levels at fc+f0 and fc−f0 are the same (refer to Eq. 27 in the supplemental material). This raises the question: what is the utility of distinguishing between the +1 and -1 orders? Is it not sufficient to distinguish between focusing on user 1 and 2.*

Response and Action Taken: Thank you very much for your insightful comments. The STC metasurface is architecturally divided into four sub-metasurfaces, each consisting of a 16×16 array of unit cells. The ability to independently control these unit cells introduces a degree of system complexity, which inherently affects the overall switching speed. This trade-off between fine-grained control and operational efficiency is a crucial factor in the design and optimization of programmable metasurfaces. To address this limitation, future improvements in control circuit performance can further enhance the switching frequency of the metasurface. In our experiment, we employed a high-resolution commercial spectrum analyzer, which is capable of differentiating between the +1 and -1 harmonic frequencies, as shown in Figure R12. Additionally, we can further refine the receiving system by employing an integrated module that combines a highly selective RF front-end with coherent demodulation in future work, enabling precise discrimination between these harmonic frequencies. First, a low-noise amplifier (LNA) and a narrowband band-pass filter are used to acquire the desired frequency range. Then, a mixer down-converts the receiving signals to baseband signals and advanced digital signal processing algorithms are subsequently applied to resolve the 2 kHz frequency offset, enabling coherent demodulation and accurate differentiation between two signals.

By leveraging multiple harmonic-frequency-dependent physical channels, the STC metasurface

enables multi-channel information transmission while simultaneously regulating the propagation direction and harmonic power distribution. This capability enhances system security by making it more difficult for illegal receivers to intercept the transmitted signals. While the STC metasurface can also achieve dual-user wireless communication by generating two spatially distinct beams without harmonic frequency-based differentiation, a purely spatial approach is inherently vulnerable to eavesdropping by receivers positioned near authorized users. In contrast, harmonic frequency-based multiplexing ensures that only legal receivers, capable of jointly decoding both the designated harmonic frequency and spatial position, can successfully reconstruct the transmitted information. **Furthermore, our system is designed to provide users with flexible configuration options, allowing them to selectively utilize harmonic frequency communication based on their security and performance requirements.** Given that the primary objective of our study is to enhance the security of the communication process, we have deliberately adopted this approach to strengthen resilience against unauthorized interception.

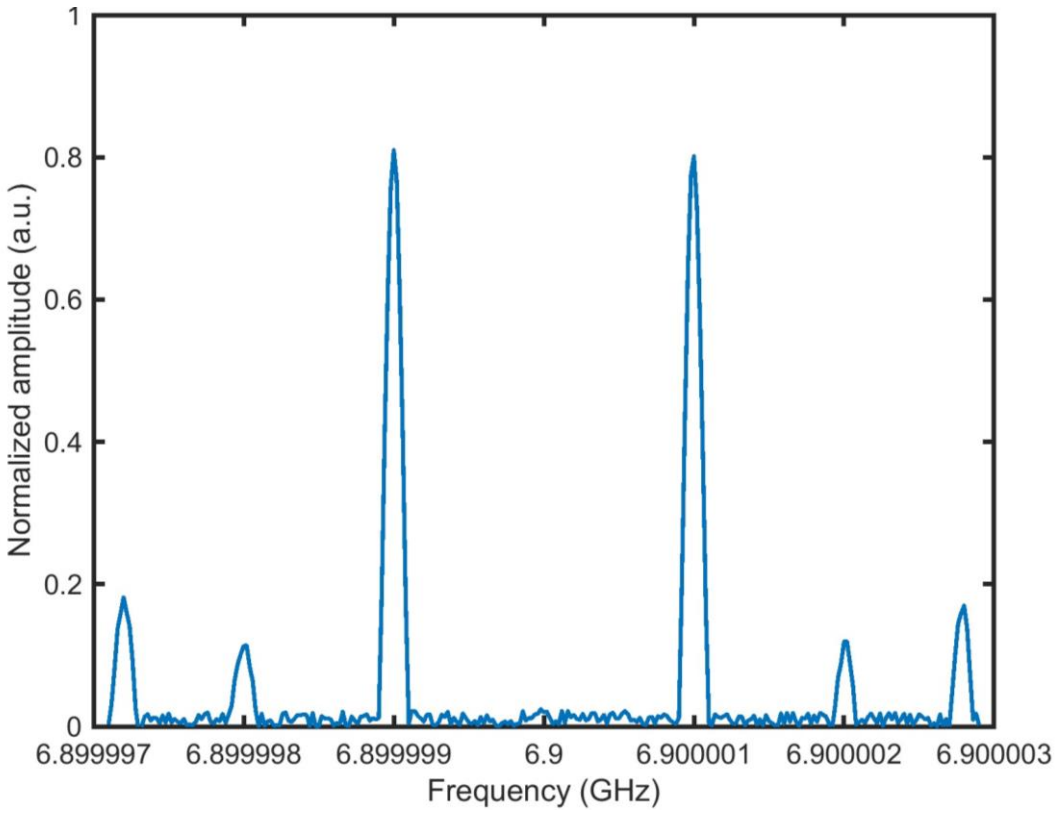

Figure R12. The measured spectral distributions for the scattered waves received from a specific direction.

Comment 3.5: *Generation and Role of HSK1 and HSK2: How are HSK1 and HSK2 generated? How are they used for decryption? Are VSK1 and VSK2 sufficient on their own? These points are unclear.*

Response and Action Taken: Thank you very much for your valuable comments. In the previous encryption scheme, HSK1 is a randomly generated 5×5 matrix composed of blocks "0" and "1" and HSK2 is then calculated according to the XOR operation. In the initial design of our system, HSK1 and HSK2 are proposed as additional security layers to complement the VSKs (VSK1 and VSK2). The original goal of HSK1 and HSK2 is to introduce an additional barrier against interception. However, in our updated design, we have streamlined the encryption mechanism to focus on VSK1 and VSK2 exclusively, recognizing that they already offer robust protection by requiring both shares for decryption. As stated in the response to ***Comment 3.1***, we acknowledge that VSK1 and VSK2 are sufficient for secure decryption and two HSKs can be used to further enhance security performance. We have simplified our approach and removed HSK1 and HSK2 from the revised version, ensuring a more streamlined and coherent encryption and decryption process. Two ciphertexts (VSK1 and VSK2) are generated using the XOR-based encryption scheme, where VSK1 is randomly generated. Each pixel of the secret information is divided into two encrypted shares (VSK1 and VSK2), ensuring that neither share contains sufficient information on its own to reveal the secret. Then, BSTCM is used to send different VSKs to two users through two harmonic frequency channels. Neither VSK1 nor VSK2 alone contains sufficient information to reconstruct the original secret information, ensuring the security of the transmission. At the receiver side, decryption is performed using a simple XOR operation between the received VSK1 and VSK2, reconstructing the original secret. We have explicitly clarifies these aspects in the revised manuscript. We sincerely appreciate your comments and believe these clarifications will improve the clarity of our work.

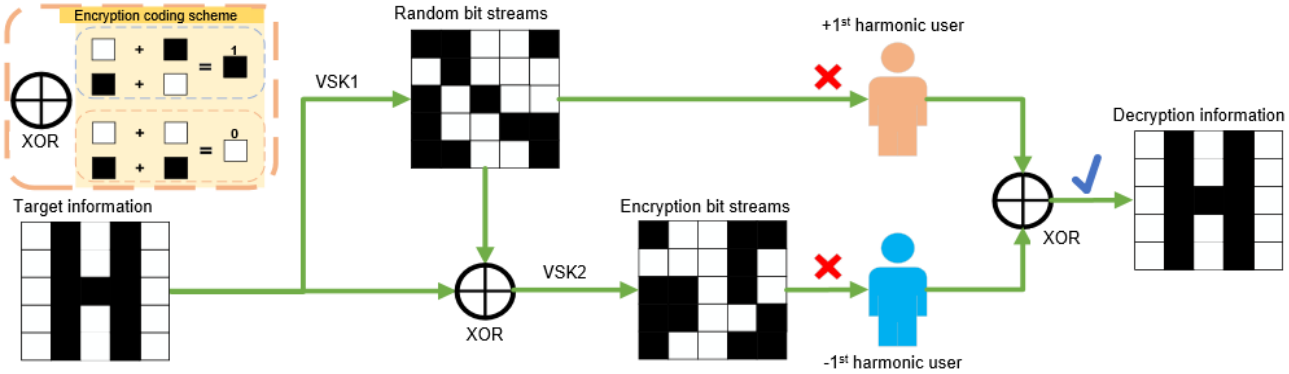

Figure R13. The encrypted encoding scheme for the secret image "H". The VSK1 is randomly generated and VSK2 is calculated by XOR operation according to the encryption coding scheme.

[Introduction]

> …Specifically, the transmitted information is encrypted into two ciphertexts, which are respectively sent to two users through two independent harmonic frequency channels constructed by the BSTCM system….

[System configuration]

> …In fact, the target information is encrypted into two secret ciphertexts. As schematically demonstrated in Fig. 1b, a legal transmitter (Alice) equipped with an EEG cap intentionally transfers two secret ciphertexts to two legal receivers (Bob and Carol) at two different harmonic frequency channels based on the encrypted communication. The eavesdropper (Eve) cannot decrypt the transmitted information unless she simultaneously obtains two ciphertexts, and knows the encryption mechanism, which is nearly impossible. Hence, the proposed system ensures high security and concealment by encrypting the secret information into two ciphertexts and transmitting these ciphertexts through two harmonic frequency channels…..

[Secure encrypted wireless communication system]

> …We present an encryption strategy designed to facilitate the secure wireless communication influenced directly by the human intention. Under this approach, the original information is recovered by XOR-based stacking of two visual shared keys (VSKs), wherein each pixel of the secret message is encrypted into two sub-pixels for two distinct users. As illustrated in Fig. 5a, VSK1 is randomly generated and VSK2 is subsequently derived through the XOR operation. Notably, neither share alone contains any discernible details regarding to the original secret, thereby ensuring that the disclosure of a single VSK does not compromise the confidentiality of the underlying message. As an illustrative example in Fig. 5a, the secret image “H”, consisting of a grid of 5×5 black pixels (digital “1”) and white pixels (digital “0”), is encrypted into two VSKs composed of 5×5 sub-pixels based on the presented encryption method, where VSK1 is randomly generated and VSK2 is subsequently derived through the XOR operation. Consequently, neither VSK1 nor VSK2 alone contains sufficient information to reveal the original secret information….

Comment 3.6: *Transmission of K1 and K2: According to Figure 25 in the supplemental material, transmitting K1 and K2 appears to require a separate transmission system. If so, why not use this system for transmitting the primary information?*

Encoding/Decoding Process: The method for encoding and decoding information using K1 and K2 remains vague and inadequately explained.

Response and Action Taken: Thank you very much for your thoughtful and constructive feedback. We emphasize that no additional transmission system is strictly required to transmit *K1* and *K2*. The proposed BSTCM system can be also used to transmit *K1* and *K2*. We acknowledge that the previous

system is more complex and we have further simplified the entire system, as stated in the response to ***Comment 3.1*** and ***Comment 3.5***. The proposed BSTCM system is designed to transmit two VSKs to two distinct users through the establishment of separate harmonic frequency channels, as shown in Figure R13. The primary purpose of our transmission strategy is to ensure that even if an illegal eavesdropper intercepts part of the signal, they cannot reconstruct the original information. The utilization of harmonic frequency channels significantly augments security by guaranteeing that accurate information can be decrypted through the combination of the received VSK1 and VSK2 in accordance with the established encryption encoding scheme, as illustrated in Figure R13. The BSTCM system integrates the transmission of VSK1 and VSK2, avoiding the need for a separate system. The existing harmonic multiplexing mechanism allows different users to receive distinct information while mitigating the risk of unauthorized access. Using the same system for secure primary data communication maintains a high level of adaptability, allowing the host user to select whether or not to incorporate harmonic-frequency-based security mechanisms. To address this concern, we have also modified Figure R14 and the relevant description to prevent any potential misinterpretations. We deeply appreciate your insightful comments, which have helped us improve the clarity and rigor of our work. We hope this response adequately addresses your concerns and demonstrates the necessity and advantage of our proposed approach.

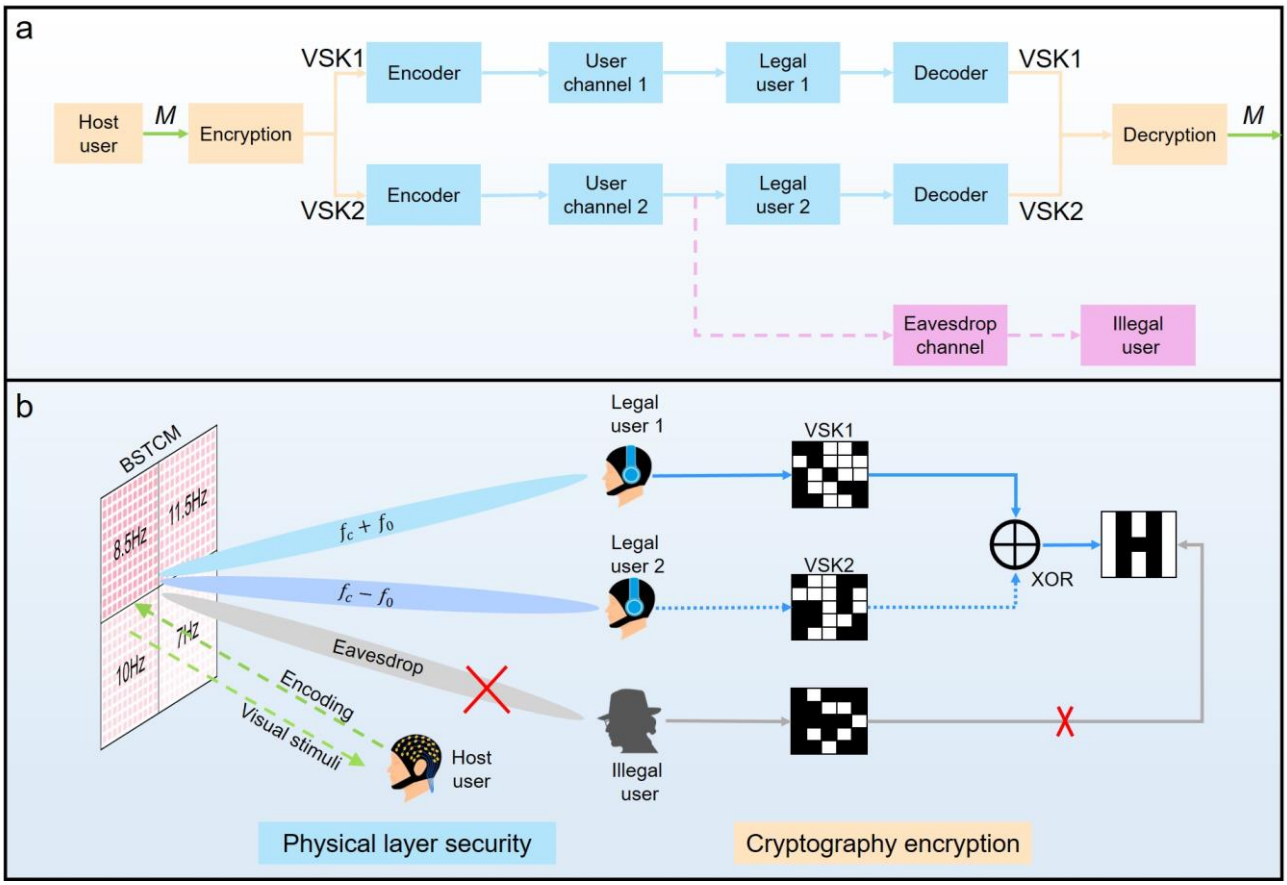

Figure R14. The proposed secure wireless communication scheme. a. The detailed flowchart of the proposed security model. b. The schematical diagram of the presented BSTCM-based secure wireless communication system.

[Supplementary Note 15: The Security Performance of the Proposed Secure Wireless Communication]

…The proposed high-security model of the BSTCM-based wireless communication system is shown in Supplementary Figure 25, which integrates cryptographic encryption method and physical layer security. The detailed secure wireless communication process is demonstrated in Supplementary Figure 25a. According to the encrypted process of Supplementary Figure 20, the target message (M) is firstly encrypted to two ciphertexts (VSK1, VSK2). The host user transmits two ciphertexts to two separate users by focusing on distinct visual stimuli, with each ciphertext traveling through dedicated harmonic channels to ensure robust security and concealment. During this communication process, two ciphertexts (VSK1, VSK2) are encoded as the spatial information and then sent to two secure user communication channels constructed from different harmonic frequencies and different spatial directions by the proposed BSTCM platform, as shown in Supplementary Figure 25b. Superimposed on the effects of channel characteristics, information propagating through the channel transforms into received by the target users. Two legal users at specific directions can receive the corresponding messages and recover the target message (VSK1, VSK2). Finally, all information from both two legal users is accurately decoded into the correct message M according to the decryption process in Supplementary Figure 20. However, the illegal user is unable to acquire any meaningful information without access to the direction and harmonic-frequency information. What's more, it is also impossible to crack the correct information from a single physical communication channel, greatly improving the security of information.….

Comment 3.7: *User Communication for Decryption: Do users need to communicate with each other to decrypt the signal? If so, how is this communication established and managed?*

Response and Action Taken: Thank you very much for raising this important question. As stated in the response to ***Comment 3.1***, no direct communication between the two users is required for successful decryption in our proposed BSTCM system. The encryption mechanism employed in our system is based on a publicly known XOR-based encoding scheme, which is detailed in the manuscript (see Figure R12). Since this encryption mechanism is fully disclosed and accessible, both users only need to receive their respective VSKs (VSK1 or VSK2) through the designated harmonic frequency channels. By utilizing distinct harmonic frequency channels, our system ensures that each user receives only their assigned data. A commercial spectrum analyzer is used to received and demodulated the transmitting data from two harmonic frequency channels, respectively. Two VSKs are acquired at the host side via post-processing. Then, two received VSKs are combined in accordance with the established XOR-based encryption scheme (illustrated in Figure R12) to accurately reconstruct the original information. Consequently, the system remains robust against potential eavesdropping, as unauthorized interception of one channel alone is insufficient to recover the original information. While our existing implementation achieves this effectively, we recognize the potential for further system optimization and improved integration according to your valuable feedback. To improve system efficiency and streamline the decryption process, two integrated modules can be placed at the user end. The integrated module will incorporate a down-conversion circuit and analog-to-digital converter to convert the received signals into digital data streams for further processing. The preprocessed data will be fed directly into an FPGA platform, where the XOR-based decryption scheme (as detailed in Figure R13) will be executed efficiently. To enhance clarity, we have revised the manuscript to explicitly emphasize that the encryption mechanism is public, and no user-to-user communication is required. We greatly appreciate the reviewer's valuable comments, which have enabled us to refine our explanations and improve the manuscript's clarity.

Comment 3.8: *While I find the results concerning the control of smart devices compelling, the same cannot be said for the secure wireless communication aspects. In addition to an incomplete and unclear description of the transmission process, there are fundamental issues that render the approach questionable:*

• The rationale for using encoding based on visual stimuli (as claimed in some parts of the manuscript) is unclear and appears unnecessary. A robust random generator would likely achieve better efficiency and simplicity.

Response and Action Taken: Thank you very much for your thoughtful and detailed comments. The necessity of visual stimulation lies in its ability to evoke frequency-specific cortical facilitation, which serves as the primary biomarker for decoding user intent in SSVEP-BCIs. Without such stimulation, the distinct, sustained oscillatory patterns required for reliable signal detection and classification would not manifest, rendering the system inoperative. The process of visual stimuli encoding is detailed in the response to ***Comment 3.1***. The primary motivation for employing visual stimuli-based encoding in our BSTCM system is to integrate multiple requirements in secure communication systems, including user interaction, real-time control, and encrypted transmission. By associating distinct visual stimuli with different encoding schemes, our system allows the host user to dynamically select which portion of the information is being transmitted. Unlike a purely random generator, our system ensures that the user's real-time EEG responses can adaptively modify the encoding pattern. Consequently, the system can respond to changes in the user's gaze or attention, which would be challenging to replicate using a conventional random generator tied to a fixed seed. In many BCI scenarios — especially those involving hands-free control or human intention recognition — the SSVEP-driven visual stimuli simultaneously serve as both a control signal and a security mechanism. This dual role helps minimize hardware complexity and eliminates the need for additional control channels. Moreover, by directing one's gaze toward specific flickering frequencies on the STC metasurface, the user intuitively manages information transmission in a way that a standard random generator cannot inherently provide. In addition, the security performance of the BSTCM system can be improved by constructing different frequency-dependent communication channels via space- and frequency-multiplexing. We have revised the manuscript to emphasize that SSVEP-based encoding is not merely a key generator but a BCI-driven mechanism for secure, interactive communication. We fully recognize that a robust random generator can offer efficiency and simplicity for cryptographic purposes. However, in the context of a BSTCM system—where user intention, information interaction, system security, and dynamic programmability are key—SSVEP-based visual stimuli provide a unique and potent solution. We trust this additional detail clarifies the necessity and advantages of our approach, and we appreciate the reviewer's feedback in helping us articulate these points more clearly.

[Design of STC metasurface]

> …Based on the encoding process, the digital symbol "0" or "1" for User 1 is encoded as the repeated data frames "M1-M0-M0-M0-M1-M0-M0-M0…" or "M1-M0-M1-M0-M1-M0-M1-M0…". Similarly, the digital symbol "0" or "1" for User 2 is encoded as repeated data frames "M2-M2-M0-M0-M2-M2-M0-M0…" or "M2-M2-M0-M0-M2-M2-M0-M0…", respectively.….

[Secure encrypted wireless communication system]

> …When the BCI operator (Alice) sequentially transmits the two bitstream sequences of two VSKs to two legal users (Bob and Carol) by the human mind, she freely gazes at the visual stimuli of the four STC metasurface regions based on the properties of the SSVEP-based BCI system. As a result, different EEG signals are generated and classified into four commands, which are then sent to the FPGA to produce the corresponding four BSTCM signals, as illustrated in Fig. 4j–m. Each gaze from the BCI operator corresponds to transmitting a single pixel of the VSK. Due to the neural response time constraints and the required high classification accuracy, the system's data transmission rate is about 0.25 bps, which can be further improved with advanced technology (refer to Supplementary Note 17). By employing the proposed BSTCM system to construct two harmonic frequency-dependent physical channels ($+1^{st}$ and -1^{st} harmonics), two VSKs (VSK1 and VSK2) are sequentially delivered to the respective legal users through these channels. During decryption, the two VSKs corresponding to each secret are XOR-based to recover the original information. The details of encryption and decryption methods can be found in Supplementary Note 12. The target image in question can be decrypted through the XOR operation of two VSKs, realizing a robust level of information security. In the future, the encryption scheme can be expanded to include more users, further enhancing the BSTCM system's security….

Comment 3.9: *The overall communication system appears overcomplicated, seemingly incorporating as many techniques as possible (e.g., ML, FPGA, STC matrix, secure communication, IoT, etc.), which results in a lack of consistency and coherence. For these reasons, I still must recommend rejecting the paper. While the communication approach shows some promise, it requires significant revision, focusing on removing unnecessary steps and providing a clearer, more streamlined methodology.*

Response: Thank you very much for your professional comment and constructive feedback. We fully understand your concerns regarding the perceived complexity of the proposed system. The core novelty of this paper is the innovative **fusion-coding method between SSVEP visual stimulation coding and metasurface space-time coding (STC) for secure information interaction at the physical layer**. In order to implement this fusion-coding mechanism, we have developed a highly integrated BSTCM system that leverages advanced technologies such as ML, FPGA-based control, STC matrix, and secure communication systems. Each module (ML for EEG recognition, FPGA for

real-time signal control, STC matrix for harmonic channels, secure encryption, etc.) addresses a specific requirement—together forming a highly integrated system that seamlessly handles both user intention (via EEG) and secure data transmission. These techniques complement each other to ensure real-time adaptation, robust security, and user-driven interaction in an integrated platform. Each incorporated technique serves a well-defined purpose:

FPGA: An FPGA is pivotal in our BSTCM system for real-time signal processing, as it fuses both visual stimuli signals and metasurface STC signals to form a BSTCM signal. This functionality is crucial to support EEG-based user interaction while simultaneously enabling encrypted information transmission. The system can quickly fuse the low-frequency visual stimuli (driving SSVEP-based EEG responses) with the higher-speed STC matrices—without interrupting or degrading either signal path by an FPGA. Consequently, the FPGA ensures seamless synchronization between user-intention cues and the information interaction process, making it indispensable for real-time adaptation and efficient management of both stimuli-driven interactions and data encryption within the BSTCM platform.

STC Matrix: The STC matrix is fundamental to our system's functionality, as it realizes the simultaneous manipulation of EM waves across both spatial and frequency domains. By precisely designing beam directions and harmonic power distributions, the STC matrix establishes independent physical channels, allowing distinct communication streams to be isolated and transmitted simultaneously. This spatiotemporal control enhances security by ensuring that unauthorized receivers must intercept signals at the correct direction and frequency, thereby mitigating eavesdropping risks. Moreover, it facilitates highly flexible, multi-user communication, supporting dynamic adaptability in various operational scenarios. In essence, the STC matrix integrates secure transmission, multi-channel support, and adaptability, serving as the cornerstone of our BSTCM system.

Machine Learning (ML): ML is indispensable for EEG signal classification, due to the inherently complex, non-stationary nature of brain signals. Traditional signal processing techniques often struggle to capture subtle temporal or frequency-domain features critical for reliable EEG interpretation. However, ML algorithms can adaptively learn from large EEG datasets, discerning patterns and fluctuations that directly correlate with user intention or cognitive state. This capability not only boosts

classification accuracy but also enables real-time responsiveness, which is vital for applications such as visual stimuli-based encoding and contactless device control. Consequently, integrating machine learning into the EEG recognition pipeline is essential to achieve robust, high-performance user interaction in the proposed BSTCM system, ensuring that dynamic brain signals can be accurately classified to corresponding control commands.

Secure Communication: The proposed secure wireless communication system integrates physical layer security with a cryptographic encryption method, ensuring that intercepted signals alone cannot reconstruct the original message. The system is implemented to enable secure information transfer between the human brain and external devices, enhancing the reliability and security of human brain's information interaction in complicated wireless environments.

IoT: Our system demonstrates contactless secure communication and remote device control, which has great potential in the field of IoT. It may pave the way for next-generation smart environments where security, scalability, and adaptability are seamlessly integrated, driving advancements in various IoT sectors. To avoid ambiguity, we have removed the relevant descriptions.

Each of these components is intricately interlinked and indispensable, forming a cohesive framework that underpins the system's overall performance. The BCI signals are first decoded into distinct control commands via ML technology. These commands are then transmitted to an FPGA, which integrates them with the STC signals and visual stimulation signals to synthesize a BSTCM signal. Finally, the BSTCM signal drives the STC metasurface to enable secure information transmission. While each technique has its specific role, they collectively address the need for real-time processing, enhanced security, and user-interactive control, which are fundamental for the proposed BSTCM system's intended use in intelligent environments. The combination of these techniques is essential to achieve the desired functionality within a unified framework. The interplay between FPGA control, STC-based harmonic multiplexing, and visual stimuli encoding ensures a highly adaptable, secure, and user-driven communication system.

Recognizing the reviewer's concerns, we have streamlined the BSTCM system and enhanced its coherence, as illustrated in Figure R14. When the BCI operator focuses on different visual stimuli

supported by the STC metasurface, distinct EEG signals are generated and classified into corresponding commands. These commands then fuse the visual stimulation signals with the STC signals, forming BSTCM signals that simultaneously enable both visual stimulation and information interaction. Based on the visual stimuli encoding mechanism, we have successfully realized a secure wireless communication system and smart device control. To address the complexity of the secure communication methodology, we have removed non-essential steps and simplified the encryption mechanism by eliminating HSK1 and HSK2, concentrating solely on VSK1 and VSK2 for secure transmission. As illustrated in Figure R12, our cryptography encryption method employs a universal XOR operation: VSK1 is randomly generated, while VSK2 is calculated via XOR based on the target information. The contents of VSK1 or VSK2 are then transmitted to two designated users via the $+1^{st}$ or -1^{st} harmonic frequency channel constructed by the STC metasurface, leveraging user gaze interaction with visual stimuli to facilitate secure and intuitive communication. We believe that these revisions have significantly improved the clarity, coherence, and overall presentation of the proposed system. We are confident that the refined manuscript now offers a streamlined methodology that effectively balances innovation with practical implementation.

Action Taken: We are grateful for the reviewer's critical insights, which have helped us identify areas for improvement. We have taken substantial steps to streamline the methodology, enhance clarity, and improve coherence while retaining the system's core innovations. We have simplified or combined overlapping functionalities to reduce complexity. We hope these revisions address the concerns raised and demonstrate the system's value as a novel and effective solution for secure communication. We sincerely thank you for your time and dedication to improving our work.

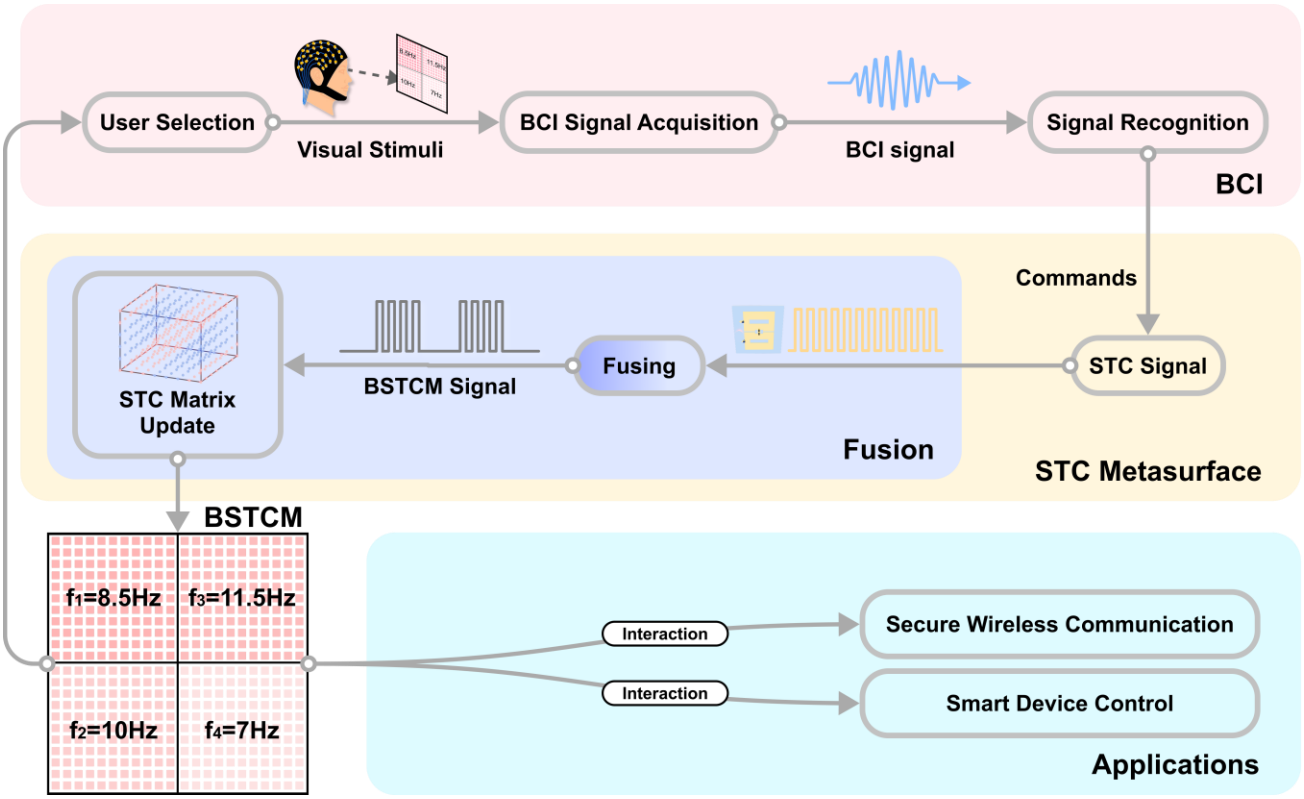

Figure R14. The detailed flow diagram of the BSTCM system. The user generates the BCI signals by gazing at visual stimuli, which are processed and classified into different commands by the machine learning method. These commands are sent to the STC metasurface, which fuses the visual stimuli signals and STC signals to form different BSTCM signals for the secure wireless communications and remote controls of smart devices through the frequency- and space-multiplexed physical channels.

[Introduction]

> …Brain-computer interface (BCI) has emerged as a cutting-edge technology in human-machine interaction and demonstrates promising applications such as brain-controlled spelling input[1,2], medical rehabilitation, and equipment control[3,4]…..
> …The SSVEP-based BCI systems utilize the brain's SSVEP response to the fixed-frequency visual stimuli for mind recognition and interaction8, offering a significant advantage of a high information transfer rate. Recently, some high-performance BCI systems have been continuously proposed[9-14], and the advancement of 6G wireless communication technology has significantly expanded the potential applications of BCIs. Hence, ensuring high information security and privacy preservation[15] for BCI users and devices becomes imperative when facilitating intelligent interactions within the constructed communication environment.….

[System configuration]

…The proposed deep fusion coding method combines the BCI visual stimulation coding with the metasurface coding for secure information interactions at the physical layer, as presented in Fig. 1a. The brain signals can be induced by low-frequency flickering visual stimuli (4~50Hz), while the STC metasurface produces a series of harmonic frequencies, revealing notable similarities in the frequency domain. The visual stimulation is crucial for eliciting the distinct oscillatory patterns necessary for user-intent decoding; and without it, the BCI system would be inoperative. Given that the metasurface's switching frequency is significantly higher than the flicker frequency, we fuse the STC signals into the high-level intervals of low-frequency visual stimuli. This integration enables the seamless fusion of the two coding mechanisms, as illustrated in Fig. 1a. To implement this fusion coding, an integrated BSTCM platform for both visual stimulation and EM-wave modulation is presented in Fig. 1, which comprises an STC metasurface and a BCI device. The metasurface element includes a meta-structure for EM-wave regulation and an LED stimulator for visual stimuli. The corresponding EEG signals can be accurately captured by a head-worn EEG cap when the user focuses on the LEDs flickering at distinct frequencies. Then, the brain's intention can be recognized to different commands through the classification process of EEG signals and the commands are transmitted to the STC metasurface, where the metasurface coding is fused with the visual stimuli to generate BSTCM control signals. This simultaneously supports visual stimulation and EM-wave modulation, opening up a wide range of potential applications.…

…The detailed flow diagram of the BSTCM system is shown in Fig. 2. The target flickering stimuli, characterized by four distinct frequencies (8.5Hz, 10Hz, 11.5Hz, and 7Hz), are supported by the STC metasurface integrated with LEDs. The interaction process begins with the user wearing an EEG cap and sitting in front of the STC metasurface. Four visual flickering stimuli with distinct frequencies are presented to the user for selection. These visual stimuli are employed to elicit specific BCI signals, while an EEG amplifier collects the BCI signals (see Supplementary Note 1 for details). Then, the BCI signals are classified into four different interaction commands, each corresponding to one of the aforementioned visual stimulation frequencies. The interaction commands are sent to the STC metasurface and the corresponding STC signals are generated. A fusion operation is conducted to fuse the STC signals and visual stimulation signals, hence generating the corresponding BSTCM signals. These BSTCM signals drive the STC metasurface to simultaneously maintain the visual stimuli of the fixed frequency and manipulate the EM waves. Finally, the BSTCM system is used to construct independent physical channels at harmonic frequencies, enabling secure wireless communication and remote control of smart devices. This integrated platform not only ensures high mobility and system compactness by reducing the external connections but also achieves real-time bidirectional interaction. This approach highlights a novel fusion of BCI and STC metasurface technologies for secure, efficient, and intelligent human-computer interaction systems. Every module of the BSTCM system architecture addresses a specific requirement yet remains inextricably linked to the others. They collectively form a highly integrated system that seamlessly handles user intention and secure data transmission, ensuring real-time adaptation, robust security, and user-driven interaction. Ultimately, the BSTCM system allows individuals to customize and modulate EM waves via cognitive intent, expanding the EM functionalities while enhancing the flexibility and interactivity in the human-machine systems.

…The STC metasurface enables precise control to the propagation direction and harmonic power distribution of the EM waves in both spatial and frequency domains. By harnessing the abundant harmonic modulation capabilities and integrating the advantages of human brain intelligence, the BSTCM system facilitates the implementation of a secure encrypted wireless communication system and enables smart device control by human mind. In the encrypted wireless communication system, the XOR-based encryption encoding method is employed to encrypt the target information[34-36]. In fact, the target information is encrypted into two secret ciphertexts. As schematically demonstrated in Fig. 1b, a legal transmitter (Alice) equipped with an EEG cap intentionally transfers two secret ciphertexts to two legal receivers (Bob and Carol) at two different harmonic frequency channels based on the encrypted communication. The eavesdropper (Eve) cannot decrypt the transmitted information unless she simultaneously obtains two ciphertexts, and knows the encryption mechanism, which are nearly impossible. Hence, the proposed system ensures high security and concealment by encrypting the secret information into two ciphertexts and transmitting these ciphertexts through two harmonic frequency channels. Ingeniously camouflaged as an LED stimulator, the STC metasurface can effectively obstruct the eavesdropper from detecting the information interaction. The proposed encrypted wireless communication system integrates the physical layer security with the cryptography encryption method, significantly enhancing the system security. Furthermore, the proposed system facilitates the remote control of smart devices, enabling users to control devices directly through the user's brain intention without the need for physical actions.

[Secure encrypted wireless communication system]

…In the proposed secure encrypted wireless communication system, the transmitted information is firstly encrypted into two ciphertexts. These ciphertexts are then securely sent to two designated users via the BSTCM system, which are seamlessly controlled by human intention. We present an encryption strategy designed to facilitate the secure wireless communication influenced directly by the human intention. Under this approach, the original information is recovered by XOR-based stacking of two visual shared keys (VSKs), wherein each pixel of the secret message is encrypted into two sub-pixels for two distinct users. As illustrated in Fig. 5a, VSK1 is randomly generated and VSK2 is subsequently derived through the XOR operation. Notably, neither share alone contains any discernible details regarding to the original secret, thereby ensuring that the disclosure of a single VSK does not compromise the confidentiality of the underlying message. As an illustrative example in Fig. 5a, the secret image "H", consisting of a grid of 5×5 black pixels (digital "1") and white pixels (digital "0"), is encrypted into two VSKs composed of 5×5 sub-pixels based on the presented encryption method, where VSK1 is randomly generated and VSK2 is subsequently derived through the XOR operation. Consequently, neither VSK1 nor VSK2 alone contains sufficient information to reveal the original secret information. When the BCI operator (Alice) sequentially transmits the two bitstream sequences of two VSKs to two legal users (Bob and Carol) by the human mind, she freely gazes at the visual stimuli of the four STC metasurface regions based on the properties of the SSVEP-based BCI system. As a result, different EEG signals are generated and classified into four commands, which are then sent to the FPGA to produce the corresponding four BSTCM signals, as illustrated in Fig. 4j–m. Each gaze from the BCI operator corresponds to transmitting a single pixel of the VSK. Due to the neural response time constraints and the required

high classification accuracy, the system's data transmission rate is about 0.25 bps, which can be further improved with advanced technology (refer to Supplementary Note 17). By employing the proposed BSTCM system to construct two harmonic frequency-dependent physical channels (+1st and -1st harmonics), two VSKs (VSK1 and VSK2) are sequentially delivered to the respective legal users through these channels. During decryption, the two VSKs corresponding to each secret are XOR-based to recover the original information. The details of encryption and decryption methods can be found in Supplementary Note 12. The target image in question can be decrypted through the XOR operation of two VSKs, realizing a robust level of information security. In the future, the encryption scheme can be expanded to include more users, further enhancing the BSTCM system's security….

[Supplementary Note 12: The Harmonic-Based Encrypted Wireless Communication Scheme]

…Firstly, the encryption coding scheme is designed based on the XOR operation. The digital "1" of the secret image can be decrypted when an exclusive OR (XOR) relationship is set between two information blocks. Whereas, the digital "0" of the secret image can be decrypted when an exclusive NOR (XNOR) relationship is set between two information blocks. The secret image can be encrypted as two information block matrices (VSK1 and VSK2). According to the encryption method, two VSKs will be sent to two different users with different harmonic frequencies, respectively. As a result, every single secret is divided into two matrices so that the effective information can be hidden as two VSK1 and VSK2. In this way, the contents of transmitting single VSKs are completely random and meaningless, which makes either VSK1 or VSK2 strongly robust against different eavesdroppers. Then, two random VSK1/VSK2 are sent to two users with two harmonic frequencies by the constructed secure encrypted wireless communication system based on the BSTCM platform….

…The secret image can be further decrypted when two VSKs are combined to the secret information according to the encryption coding scheme….

…It is clear that the extracted results of visual images are consistent well with the theoretical results, the corresponding secret pixel can be deciphered by the XOR operation of two shared pixels (VSK1 and VSK2), each of which is composed of 5 × 5 shared secret bit streams….

…In order to further validate the proposed encryption wireless communication system, we also demonstrate the encryption process for three additional letter patterns "E", "L", and "I", as illustrated in Supplementary Figure 22. According to the encrypted encoding scheme in Supplementary Figure 20, the grayscale matrices of the three letters, consisting of digital 0 and 1, are encrypted into two VSKs, where VSK1 is the same and VSK2 is respectively calculated. In addition, VSK1 of three letters are randomly generated and there corresponding VSK2 are calculated. Finally, all VSKs are further encoded into random transmitting streams, which cannot disclose the content of transmitting information….

[Supplementary Note 15: Security Performance of the Proposed Secure Wireless Communication]

…The proposed high-security model of the BSTCM-based wireless communication system is shown in Supplementary Figure 25, which integrates cryptographic encryption method and physical layer security. The detailed secure wireless communication process is demonstrated in Supplementary Figure 25a. According to the encrypted process of Supplementary Figure 20, the target message (M) is firstly encrypted to two ciphertexts (VSK1, VSK2). The host user transmits two ciphertexts to two separate users by focusing on distinct visual stimuli, with each ciphertext traveling through dedicated harmonic channels to ensure robust security and concealment. During this communication process, two ciphertexts (VSK1, VSK2) are encoded as the spatial information and then sent to two secure user communication channels constructed from different harmonic frequencies and different spatial directions by the proposed BSTCM platform, as shown in Supplementary Figure 25b. Superimposed on the effects of channel characteristics, information propagating through the channel transforms into received by the target users. Two legal users at specific directions can receive the corresponding messages and recover the target message (VSK1, VSK2). Finally, all information from both two legal users is accurately decoded into the correct message M according to the decryption process in Supplementary Figure 20. However, the illegal user is unable to acquire any meaningful information without access to the direction and harmonic-frequency information. What's more, it is also impossible to crack the correct information from a single physical communication channel, greatly improving the security of information….
... We have presented a method integrating the physical layer security and the cryptography encryption as described above….