

Toward Efficient and Privacy-Aware eHealth Systems: An Integrated Sensing, Computing, and Semantic Communication Approach

Yinchao Yang, Yahao Ding, Zhaohui Yang, Chongwen Huang, Zhaoyang Zhang, Dusit Niyato, *Fellow, IEEE*, and Mohammad Shikh-Bahaei, *Senior Member, IEEE*

Abstract—Real-time and contactless monitoring of vital signs, such as respiration and heartbeat, alongside reliable communication, is essential for modern healthcare systems, especially in remote and privacy-sensitive environments. Traditional wireless communication and sensing networks fall short in meeting all the stringent demands of eHealth, including accurate sensing, high data efficiency, and privacy preservation. To overcome the challenges, we propose a novel integrated sensing, computing, and semantic communication (ISCSC) framework. In the proposed system, a service robot utilises radar to detect patient positions and monitor their vital signs, while sending updates to the medical devices. Instead of transmitting raw physiological information, the robot computes and communicates semantically extracted health features to medical devices. This semantic processing improves data throughput and preserves the clinical relevance of the messages, while enhancing data privacy by avoiding the transmission of sensitive data. Leveraging the estimated patient locations, the robot employs an interacting multiple model (IMM) filter to actively track patient motion, thereby enabling robust beam steering for continuous and reliable monitoring. We then propose a joint optimisation of the beamforming matrices and the semantic extraction ratio, subject to computing capability and power budget constraints, with the objective of maximising both the semantic secrecy rate and sensing accuracy. Simulation results validate that the ISCSC framework achieves superior sensing accuracy, improved semantic transmission efficiency, and enhanced privacy preservation compared to conventional joint sensing and communication methods.

Index Terms—Integrated sensing and communication, semantic communication, and vital sign detection.

I. INTRODUCTION

CONTINUOUS and real-time monitoring of vital signs, such as respiration and heartbeat, plays a pivotal role in the early detection and prevention of potentially life-threatening conditions [1]. Timely identification of abnormalities like irregular heartbeats is essential, as they often serve as early indicators of cardiovascular disorders [2]. Detecting these irregularities at an early stage enables prompt medical

intervention, reducing the risk of severe complications and improving patient outcomes.

Vital sign monitoring methods can be broadly classified into two categories: contact-based and contactless. Traditional contact-based approaches, such as smartwatches and multi-parameter monitors, require direct physical contact with the body, which can cause discomfort and inconvenience for patients [3]. In contrast, contactless vital sign detection has gained considerable attention due to its ability to provide non-intrusive monitoring while preserving patient privacy [4]. These systems utilise wireless signals, such as millimetre-wave (mmWave), to detect subtle physiological movements, including respiration, heartbeat, and other body motions, by analysing signal reflections from the human body [5], [6]. The integration of contactless monitoring into eHealth systems offers significant advantages for remote healthcare, telemedicine, and personalised health tracking, positioning contactless sensing as a key driver in the evolution of next-generation intelligent healthcare solutions [7].

Seamless data exchange, both within and between healthcare facilities, is also fundamental to the effectiveness of eHealth systems. Specifically, the reliable transmission of data during continuous sensing operations is paramount for delivering effective and responsive healthcare [8]. Integrated sensing and communication (ISAC) presents a promising paradigm by unifying wireless communication and contactless sensing functionalities within a single system [9]–[12]. This dual capability allows systems to concurrently support both communication and sensing tasks, thereby enhancing system efficiency and spectrum utilisation [13]–[21].

While ISAC optimises spectrum utilisation by enabling dual functionality in sensing and communication, next-generation healthcare systems demand more intelligent and efficient data transmission mechanisms [22]–[24]. Conventional wireless communication frameworks typically emphasise the reliable delivery of raw data, which can lead to excessive bandwidth consumption and the transmission of redundant or non-informative content. In contrast, modern eHealth applications require semantically aware communication strategies that prioritise clinically meaningful information, thereby enabling timely and accurate medical decision-making [8]. Semantic communication addresses this need by shifting the focus from bit-level accuracy to the transmission of relevant semantic content [25]–[27]. By leveraging knowledge bases (KBs), such as electronic health records (EHRs) which contain patient

Yinchao Yang, Yahao Ding, and Mohammad Shikh-Bahaei are with the Department of Engineering, King's College London, London, UK. (emails: yinchao.yang@kcl.ac.uk; yahao.ding@kcl.ac.uk; m.sbahaei@kcl.ac.uk).

Dusit Niyato is with the College of Computing and Data Science, Nanyang Technological University, Singapore 639798, Singapore (email: dnyato@ntu.edu.sg).

Zhaohui Yang, Chongwen Huang and Zhaoyang Zhang are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China, and Zhejiang Provincial Key Lab of Information Processing, Communication and Networking (IPCAN), Hangzhou, Zhejiang, 310007, China. (emails: yang_zhaohui@zju.edu.cn; chongwenhuang@zju.edu.cn; ning_ming@zju.edu.cn).

medical records, semantic communication enables medical devices to convey high-level meanings instead of transmitting raw physiology or clinical data [28]. For example, instead of transmitting the exact message “Patient heart rate increased from 100 BPM to 150 BPM,” a semantically aware system may send “Patient heart rate increased sharply,” thereby reducing data overhead while preserving essential clinical meaning. The exact data are stored in EHRs and shared across medical devices and, potentially, interconnected healthcare facilities. This provides a complete view of a patient’s health over time, helping to create a more efficient and responsive healthcare system [29]. Semantic communication has attracted considerable research attention in recent years. For instance, the authors in [30] proposed a machine learning-based semantic text communication system that is robust against physical channel noise and literal semantic impairments, such as misinterpretation of contextual meaning and word ambiguity. Their results demonstrated enhanced interpolation at the receiver with reduced inference time. In [31], a hybrid semantic-Shannon communication system was introduced, allowing each subcarrier to dynamically select between semantic and Shannon communication modes to optimise both delay and accuracy requirements. Furthermore, the authors in [32] explored the use of relays as edge servers to deliver semantic communication services for both semantic users with abundant computational resources and conventional users with limited capabilities. Their results showed improved communication performance while minimising the use of resources such as bandwidth and transmit power. Nevertheless, despite these notable advancements, current research predominantly targets semantic communication enhancements in general wireless networks, while the exploration of semantic communication applications within the eHealth domain remains relatively limited [33]–[35].

Although ISAC and semantic communication each demonstrate considerable potential individually, there remains a critical gap in research exploring their integration for eHealth. In an eHealth setting, joint sensing and semantic communication can enhance both the accuracy of patient monitoring and the efficiency of medical data transmission [28]. For instance, radar sensors can detect subtle physiological movements [36], while semantic processing can intelligently extract and transmit only the most clinically relevant data [37], minimising network congestion and enabling faster medical decision-making. Another significant advantage of integrating ISAC and semantic communication is the enhancement of privacy and security for patient data. Privacy and security are especially critical in healthcare systems due to the sensitive nature of medical information. Although several recent works [38]–[41] have explored privacy-preserving mechanisms for eHealth, those works overlook the potential of semantic-level protection. Medical information, including health records, diagnostic images, and treatment plans, is highly sensitive to unauthorised access. In ISAC systems, such data may be exposed to eavesdropping [7]. While precoding techniques can enhance physical layer security, their effectiveness typically relies on non-line-of-sight (NLoS) conditions, or wiretap channel models, where the eavesdropper experiences degraded channel quality.

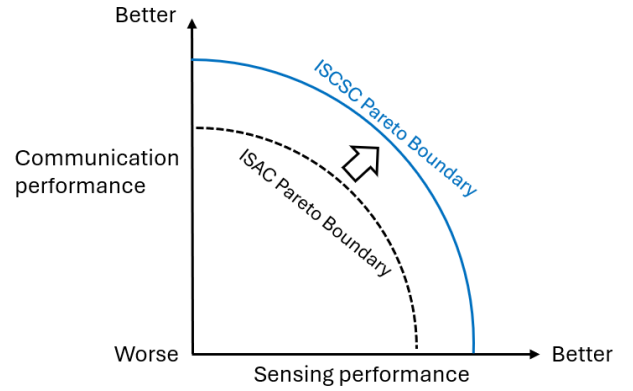


Figure 1: An illustration of the performance Pareto boundary of ISAC and ISCSC systems.

Semantic communication further reinforces data protection by eliminating the transmission of raw physiological signals. Instead, successful decoding depends on access to shared KBs. When KBs are managed in compliance with regulations such as the general data protection regulation (GDPR), unauthorised access is effectively prevented [42]. Consequently, even if semantic information is intercepted during transmission, eavesdroppers are unable to reconstruct meaningful content without the corresponding KBs.

To bridge the research gap concerning the integration of ISAC and semantic communication in healthcare systems, we propose a novel framework that effectively integrates these technologies to enable advanced healthcare applications. We refer to the proposed framework as **integrated sensing, computing, and semantic communication (ISCSC)**. Through sensing-assisted semantic communication, the ISCSC framework utilises environmental sensing to intelligently select and prioritise medical data for efficient transmission. This method not only significantly enhances data rates but also safeguards patient privacy by preventing data transmission to unintended recipients. Conversely, semantic communication-assisted sensing facilitates real-time acquisition and transmission of vital signs and medical information, enabling timely anomaly detection and further reinforcing data privacy by eliminating the need to transmit raw medical data. A comparison of ISCSC and ISAC performance in sensing and communication is illustrated in Fig. 1.

In summary, the key contributions of this paper are as follows:

- 1) We propose a novel ISCSC framework specifically designed for eHealth applications, with an emphasis on vital sign detection. The proposed ISCSC design enhances data rate and privacy while preserving the same sensing accuracy as the ISAC design.
- 2) We developed a privacy-aware predictive beamforming strategy to enable simultaneous downlink semantic communication to medical devices and sensing of patients. This strategy ensures that communication signals remain confidential and are not accessible to unintended recipients (e.g., patients being sensed).

Table I: List of main symbols.

Symbol	Description
K	Number of medical devices
$\mathcal{K}$	Set of medical devices
L	Number of patients
$\mathcal{L}$	Set of patients
$\mathbf{x}(t)$	Transmitted signal from the service robot
$\mathbf{c}(t)$	Communication signal with semantic information
$\mathbf{z}(t)$	Sensing signal
$\mathbf{w}_k$	Transmit beamforming vector of medical device k
$\mathbf{W}$	Transmit beamforming matrix of all medical devices
$\mathbf{r}_l$	Transmit beamforming vector of patient l
$\mathbf{R}$	Transmit beamforming matrix of all patients
$\mathbf{R}_x$	Covariance matrix of the transmit signal
$y_k(t)$	Received signal of medical device k
$\mathbf{h}_k$	Channel between the robot and medical device k
$y_l(t)$	Received signal of patient l
$\mathbf{h}_l$	Channel between the robot and patient l
σ_c^2, σ_r^2	Communication noise
$f_{r,l}, f_{h,l}$	Respiration and heartbeat frequencies of patient l
$A_{r,l}, A_{h,l}$	Respiration and heartbeat amplitudes of patient l
$m_{r,l}(t)$	Respiration signal of patient l
$m_{h,l}(t)$	Heartbeat signal of patient l
$\hat{\mathbf{y}}_l(t)$	Echo signal
$\phi(t)$	Instantaneous phase of the echo signal
$\mathbf{q}_{l,t}$	State vector of a patient
$\mathbf{g}_1(\cdot), \mathbf{g}_2(\cdot)$	State transition and Measurement functions
$\mathbf{u}_t, \mathbf{z}_t$	State transition and Measurement noise vectors
ρ_k	Semantic extraction ratio of medical device k
P_{comp}	Computing power required for semantic extraction
P_{cs}	Signal transmission power

- 3) To ensure accurate sensing in dynamic environments, the system employs an interacting multiple model (IMM) filter for real-time prediction of patient motion [43]. This enables dynamic beam steering to track patient positions, significantly improving the reliability and continuity of vital sign monitoring.

The remainder of this paper is organised as follows. Section II presents the system model of the ISCSC framework. Section III details the patient motion models and the associated tracking techniques. Section IV defines the performance metrics that are used to evaluate the ISCSC system. Section V formulates the optimisation problem and introduces the proposed algorithms. Section VI provides simulation results and performance analysis, and Section VII concludes the paper. The main symbols used in this paper are summarised in Table I.

List of Notations:

Capital boldface letters denote matrices, while lowercase boldface letters denote vectors. Scalars are represented by standard lowercase or uppercase letters. The set of complex numbers is denoted by $\mathbb{C}$, with $\mathbb{C}^{n \times 1}$ and $\mathbb{C}^{m \times n}$ representing an n -dimensional complex vector and an $m \times n$ complex matrix, respectively. The identity matrix is denoted by $\mathbf{I}$, and the all-zero matrix by $\mathbf{0}$. The superscripts $(\cdot)^T$ and $(\cdot)^H$ represent the transpose and Hermitian (conjugate transpose), respectively. The trace and rank of a matrix are denoted by $\text{Tr}(\cdot)$ and $\text{rank}(\cdot)$, respectively. The expectation operator is denoted by $\mathbb{E}[\cdot]$, and $|\cdot|$ represents the absolute value or magnitude. The notation $\succeq$ indicates positive semi-definiteness. The operators $\Re\{\cdot\}$ and $\Im\{\cdot\}$ extract the real and imaginary

parts, respectively. The complex Gaussian distribution with zero mean and variance σ^2 is denoted by $\mathcal{CN}(0, \sigma^2)$.

II. SYSTEM MODEL

As illustrated in Fig. 2, we consider the design of the ISCSC system within a hospital environment, comprising a service robot, multiple patients, and various medical devices. The service robot is equipped with a uniform linear array (ULA) consisting of N antennas and serves as a dual-functional node for both downlink communication and patient sensing. Let $\mathcal{K} = \{1, \dots, K\}$ and $\mathcal{L} = \{1, \dots, L\}$ denote the sets of medical devices and patients, respectively. Specifically, the robot establishes wireless communication links with K medical devices, where each device $k \in \mathcal{K}$ is equipped with a single antenna. Concurrently, the robot performs vital sign detection for L patients, with each patient $l \in \mathcal{L}$ modelled as a point target. During the concurrent operation of communication and sensing functionalities, ensuring the confidentiality of transmitted information is critical. Specifically, communication signals directed toward medical devices, often containing sensitive patient data, must be secured against unauthorised access, including from the individuals being sensed. The key procedures of the ISCSC eHealth system are summarised as follows:

- 1) Data Collection: The service robot acquires sensing data from the patients using sensors like mmWave radar.
- 2) Data Processing: The acquired data are processed to extract patient locations and vital physiological parameters.
- 3) Computing and Decision Making: Semantic information is derived based on the processed sensing data to support treatments.
- 4) Integrated Signal Transmission: A unified signal comprising both sensing and semantic information is transmitted for continuous monitoring and further communication.

Fig. 3 illustrates the key modules and their benefits within the proposed system. Each component is discussed in detail in the subsequent sections.

A. Signal Model

In the considered ISCSC system, the robot simultaneously transmits both communication and sensing signals, with beamformers assisting in accurately directing the signals toward the desired direction (i.e., the chest of a patient). Precise sensing beamforming is particularly crucial for accurately analysing heart rate and respiratory rate from the sensing signal, ensuring that each patient's vital sign is properly captured. This highlights the essential role of beamformers in the system. As patients are in motion, the robot dynamically predicts their positions in each time slot and steers the beams accordingly to maintain precise tracking and optimal signal reception. Consequently, the transmitted signal from the robot can be expressed as [44]:

$$\mathbf{x}(t) = \mathbf{W}\mathbf{c}(t) + \mathbf{R}\mathbf{z}(t), \quad (1)$$

where $\mathbf{W} \in \mathbb{C}^{N \times K}$ and $\mathbf{R} \in \mathbb{C}^{N \times L}$ are the communication and sensing beamforming matrices, respectively. Note that the design of $\mathbf{R}$ is based on the predicted states (e.g., angle) of

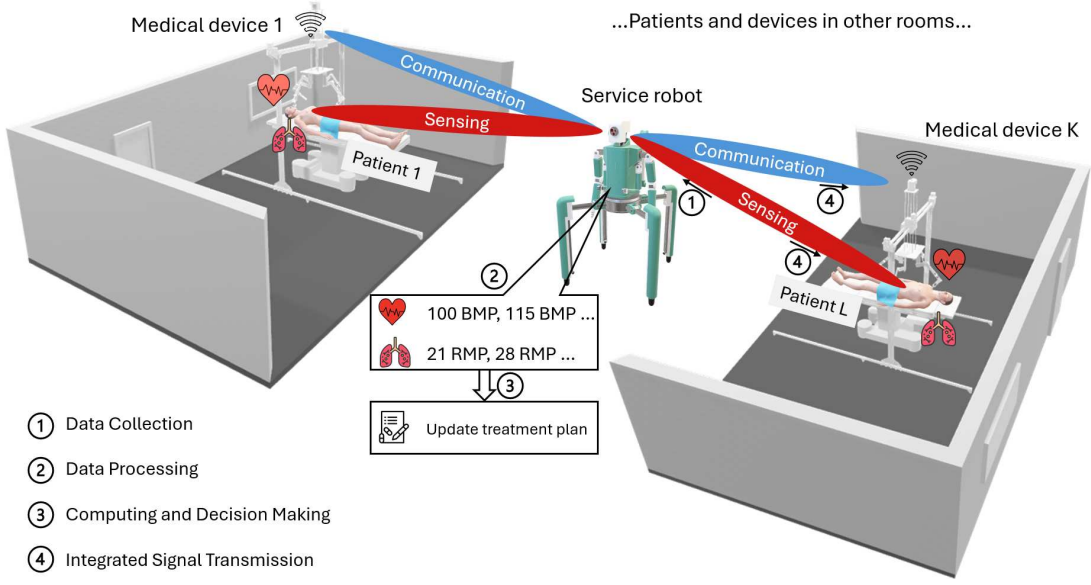


Figure 2: An ISCSC-enabled eHealth system, where a service robot engages in semantic communication with multiple medical devices and performs sensing of multiple patients to determine their locations and vital signs.

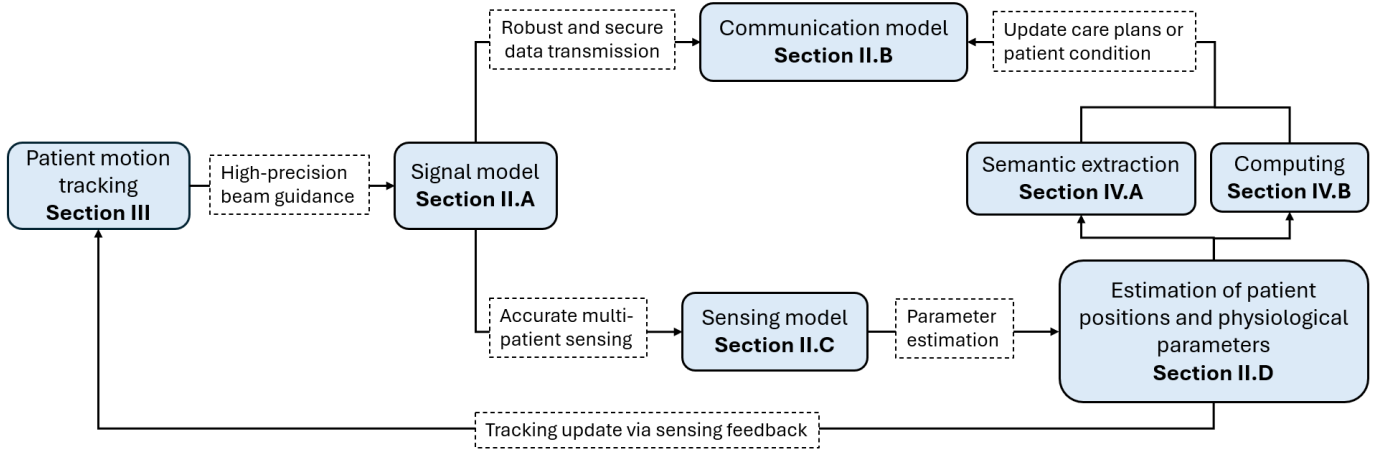


Figure 3: Flowchart of the proposed ISCSC eHealth system, illustrating the key components.

patients. The semantic message is denoted by $\mathbf{c} \in \mathbb{C}^{K \times 1}$, and $\mathbf{z} \in \mathbb{C}^{L \times 1}$ is the sensing signal. The semantic message $c_k \in \mathbf{c}$ is generated from the conventional communication message m_k , i.e., $c_k = f(m_k)$. Here, $f(\cdot)$ is a function, such as Transformers, that maps a conventional communication message to a semantic message. As an example, the original message m_k “Patient IDxx01’s heart rate is 80 BPM and respiration rate is 15 RMP; Patient IDxx02’s heart rate is 140 BPM and respiration rate is 30 RMP” can be abstracted into a high-level semantic message c_k “Patient IDxx01 is stable; Patient IDxx02 needs attention,” with higher priority assigned to the critical status of Patient IDxx02. The detailed physiological data are stored in KBs (such as the EHRs) and are accessible to authorised medical devices.

Without loss of generality, the following three assumptions are made to simplify the system design [45]:

- There is no correlation between the confidential message

and the radar signal, i.e., $\mathbb{E}(\mathbf{c}\mathbf{z}^H) = \mathbf{0}_{K \times L}$.

- There is no correlation between the confidential messages for different medical devices, i.e., $\mathbb{E}(\mathbf{c}\mathbf{c}^H) = \mathbf{I}_K$.
- There is no correlation between the sensing signals for different patients, i.e., $\mathbb{E}(\mathbf{z}\mathbf{z}^H) = \mathbf{I}_L$.

As a result, the covariance matrix of the transmit waveform is given by

$$\mathbf{R}_x = \mathbb{E}[\mathbf{x}\mathbf{x}^H] = \sum_{k=1}^K \mathbf{W}_k + \sum_{l=1}^L \mathbf{R}_l, \quad (2)$$

where $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$ and $\mathbf{R}_l = \mathbf{r}_l \mathbf{r}_l^H$. Here, $\mathbf{w}_k \in \mathbb{C}^{N \times 1}$ and $\mathbf{r}_l \in \mathbb{C}^{N \times 1}$ represent the transmit beamforming vectors for the k -th medical device and the l -th patient, respectively.

B. Communication Model

Once the joint signal is transmitted by the service robot, the received signal at the k -th medical device can be expressed as

[46]:

$$y_k(t) = \mathbf{h}_k^H \mathbf{x}(t) + n_k, \quad (3)$$

where $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ is the channel vector for the k -th medical device, $\mathbf{x}(t) \in \mathbb{C}^{N \times 1}$ is the transmitted signal, and $n_k \sim \mathcal{CN}(0, \sigma_c^2)$ is the communication noise for the k -th medical device. On the patient side, the received signal can be formulated by

$$y_l(t) = \alpha_l \mathbf{a}^H(\theta_l) \mathbf{x}(t) + n_l = \mathbf{h}_l^H \mathbf{x}(t) + n_l, \quad (4)$$

where α_l is the path-loss coefficient for the l -th patient, $n_l \sim \mathcal{CN}(0, \sigma_r^2)$ is the sensing noise for the l -th patient. The steering vector is denoted by $\mathbf{a}(\theta_l) \in \mathbb{C}^{N \times 1}$ with θ_l being the angle of the chest of l -th patient.

With (3), we can derive the signal-to-interference-plus-noise ratio (SINR) for the k -th medical device. It is expressed as:

$$\gamma_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\left| \mathbf{h}_k^H \sum_{k'=1, k' \neq k}^K \mathbf{w}_{k'} \right|^2 + \left| \mathbf{h}_k^H \sum_{l=1}^L \mathbf{r}_l \right|^2 + \sigma_c^2}. \quad (5)$$

Similarly, with (4), the SINR for the l -th patient is given by

$$\Gamma_{l|k} = \frac{|\mathbf{h}_l^H \mathbf{w}_k|^2}{\left| \mathbf{h}_l^H \sum_{k'=1, k' \neq k}^K \mathbf{w}_{k'} \right|^2 + \left| \mathbf{h}_l^H \sum_{l=1}^L \mathbf{r}_l \right|^2 + \sigma_r^2}. \quad (6)$$

C. Sensing Model

The echo signal received by the service robot of the l -th patient is given by:

$$\hat{\mathbf{y}}_l(t) = \beta_l e^{j\phi_l(t-\tau)} \mathbf{a}(\theta_l) \mathbf{a}^H(\theta_l) \mathbf{x}(t-\tau) + \mathbf{n}_l, \quad (7)$$

where β_l is the round-trip path loss coefficient and $\mathbf{n}_l \in \mathbb{C}^{N \times 1}$ is noise vector with zero mean and variance of $\sigma_e^2 \mathbf{I}$. In addition, $\phi_l(t-\tau)$ captures the effect of chest motion together with the propagation delay τ . It can be formulated as:

$$\phi_l(t-\tau) = \frac{4\pi}{\lambda} (m_l(t-\tau)), \quad (8)$$

with λ denoting the signal wavelength. The time delay τ is determined by the distance between the service robot and the patient, such as $\tau = \frac{2d}{c}$, with d being the patient-robot distance, and c representing the speed of light. Additionally, $m_l(t)$ represents the chest motion of patient l .

For each patient, the chest motion $m_l(t)$ is primarily attributed to respiration and heartbeat, and can therefore be modelled as [5], [47], [48]:

$$m_l(t) = m_{0,l} + m_{r,l}(t) + m_{h,l}(t), \quad (9)$$

where $m_{0,l}$ denotes body motion considered as interference (e.g., rolling in bed, moving in a wheelchair, or walking), $m_{r,l}(t)$ represents displacement due to respiration, and $m_{h,l}(t)$ represents displacement due to heartbeat. The respiration and heartbeat-induced displacements can be modelled by the sinusoidal functions:

$$m_{r,l}(t) = A_{r,l} \sin(2\pi f_{r,l} t), \quad m_{h,l}(t) = A_{h,l} \sin(2\pi f_{h,l} t), \quad (10)$$

where $A_{r,l}$ and $A_{h,l}$ denote the amplitudes of respiration and heartbeat, respectively, measured in millimetres, and $f_{r,l}$ and

$f_{h,l}$ represent their corresponding frequencies, measured in Hertz.

Remark 1. Equation (9) can be generalised to represent a variety of biological signals, including those related to sleep monitoring and blood pressure.

D. Estimation of Patient Position and Physiological Parameters

From the received echo, the service robot estimates the angle, heart rate, and respiration rate of each patient, i.e., the parameters embedded in (7). The first step is to estimate the time delay $\tilde{\tau}$, which provides the distance information between the robot and the patient. This is achieved using a matched filter that correlates the received signal with a time-shifted and conjugated version of the transmitted signal. According to [49], the time delay is estimated as:

$$\tilde{\tau} = \arg \max_{\tau} \left| \int_0^{\Delta T} \hat{\mathbf{y}}_l(t) \mathbf{x}^*(t-\tau) dt \right|^2, \quad (11)$$

where $\mathbf{x}^*(t-\tau)$ represents the conjugate of the transmitted signal with a time shift τ . The integration occurs over the observation interval $[0, \Delta T]$. With the estimated time delay, we calculate the instantaneous phase from the delay-compensated echo signal as follows:

$$\hat{\phi}_l(t) = \frac{1}{N} \sum \tan^{-1} \left(\frac{\Im\{\hat{\mathbf{y}}_l(t)\}}{\Re\{\hat{\mathbf{y}}_l(t)\}} \right), \quad (12)$$

where the calculation averages the phase across all channels, improving robustness against noise and ensuring accurate extraction of respiration and heartbeat components.

With the estimated phase shift $\hat{\phi}_l(t)$, both respiration and heartbeat components are inherently embedded, resulting in a composite signal. To effectively separate these physiological signatures, we employ the variational mode decomposition (VMD) method, which decomposes $\hat{\phi}_l(t)$ into a finite set of intrinsic mode functions (IMFs) [50]. This can be expressed as $\hat{\phi}_l(t) = \sum_{u=1}^U \hat{\phi}_{l,u}(t)$, where U denotes the total number of IMFs. Given the distinct spectral characteristics of respiration and heartbeat signals, they are primarily concentrated in different IMFs. Accordingly, we denote the respiration-related and the heartbeat-related components as $\hat{\phi}_{RR} = \sum_{i \in \mathcal{U}} \hat{\phi}_i(t)$ and $\hat{\phi}_{HR} = \sum_{j \in \mathcal{U}} \hat{\phi}_j(t)$, respectively, with $\mathcal{U} = \{1, \dots, U\}$. The summation reflects the fact that physiological signals may be distributed across multiple IMFs. Once the relevant IMFs are identified, we proceed to estimate the respiration and heartbeat rates using a frequency-domain approach based on the Fourier transform (FT). Although more advanced algorithms exist, such as in [51], the FT method provides a straightforward and effective way to analyse frequency components. The process involves four key steps: bandpass filtering, Fourier transform, peak detection, and amplitude estimation. We describe each step below.

- 1) To isolate the respiration and heartbeat components from the estimated phase, we apply bandpass filters tailored to their respective frequency ranges. The filters are denoted by $\text{BPF}(\cdot)$.

- A bandpass filter with a frequency range of 0.1–0.5 Hz is applied to isolate the low-frequency component associated with human respiration. This range corresponds to typical breathing rates of approximately 6–30 respirations per minute (RPM).

$$\hat{m}_{r,l}(t) = \text{BPF}(\hat{\phi}_{\text{RR}}(t), 0.1, 0.5). \quad (13)$$

- Subsequently, another bandpass filter with a frequency range of 1–2.5 Hz is applied to extract the high-frequency component associated with cardiac activity. This range corresponds to typical human heart rates of approximately 60–150 beats per minute (BPM).

$$\hat{m}_{h,l}(t) = \text{BPF}(\hat{\phi}_{\text{HR}}(t), 1, 2.5). \quad (14)$$

- 2) After filtering, we perform an FT on the respiration and heartbeat components to convert the signals from the time domain to the frequency domain. We use $\mathcal{F}[\cdot]$ to denote the FT. The resulting respiration and heartbeat spectra are given by, respectively:

$$M_{r,l}(\omega) = \mathcal{F}[\hat{m}_{r,l}(t)], \quad M_{h,l}(\omega) = \mathcal{F}[\hat{m}_{h,l}(t)]. \quad (15)$$

- 3) We then identify the dominant peaks in the respiration and heartbeat spectra. The respiration rate and heartbeat rate are determined from these peaks using:

$$\hat{f}_{r,l} = \arg \max_{\omega \in (0.1, 0.5)} |M_{r,l}(\omega)|, \quad \hat{f}_{h,l} = \arg \max_{\omega \in (1, 2.5)} |M_{h,l}(\omega)|. \quad (16)$$

Thus, the heartbeat and respiration rates can be estimated as $60 \times \hat{f}_{h,l}$ BPM and $60 \times \hat{f}_{r,l}$ RPM, respectively.

- 4) The amplitudes of the respiration and heartbeat components are estimated by identifying the maximum values of their magnitude spectra within the respective frequency bands. The respiration and heartbeat amplitudes are given by:

$$\hat{A}_{r,l} = \max_{\omega \in (0.1, 0.5)} |M_{r,l}(\omega)|, \quad \hat{A}_{h,l} = \max_{\omega \in (1, 2.5)} |M_{h,l}(\omega)|. \quad (17)$$

Finally, we can proceed with angle estimation. By applying high-resolution angle estimation techniques, such as multiple signal classification (MUSIC) [52] or estimation of signal parameters via rotational invariance (ESPRIT) [53], the angle of arrival (AoA) can be estimated. By leveraging these signal processing techniques, the service robot can accurately determine the patient's location and monitor vital signs in a contactless manner.

III. PATIENT MOTION TRACKING VIA IMM FILTER

The accuracy of the vital sign estimation as detailed in Section. II.D is fundamentally dependent on the quality of the received echo signal. To ensure a strong echo signal, the sensing beams must be precisely and continuously aimed at the patient's chest. However, since patients are dynamic, a static beam is insufficient. Therefore, a critical component of the proposed framework is the ability to track and predict patient motion in real-time. This section presents the IMM filter used to accomplish this, which enables the predictive beamforming essential for robust and reliable sensing.

A. State Evolution Model

To accurately track patient positions, their kinematic behaviour must be properly modelled. In this paper, we assume that during each time slot, a patient can move in one of four possible directions: left, right, up, down, with constant velocity, or remain stationary. As an example, these movements correspond to rolling on the bed (left/right), transitioning to a seated position (up), lying down from a seated position (down), or remaining still (stationary). Alternatively, these movements can be a patient on a wheelchair moving in one of the four directions, or remaining steady. For analytical tractability, we consider horizontal and vertical movements to be parallel to the antenna array, neglecting any minor rotational or out-of-plane components.

For each patient l at time slot t , the state model for left(+) or right(-) movement is formulated as follows [54]:

$$\begin{aligned} \theta_{l,t} &= \theta_{l,t-1} \pm d_{l,t-1}^{-1} v_{l,t-1} \Delta T \sin \theta_{l,t-1} + u_{\theta}, \\ d_{l,t} &= d_{l,t-1} \mp v_{l,t-1} \Delta T \cos \theta_{l,t-1} + u_d, \\ v_{l,t} &= v_{l,t-1} + u_v, \\ \beta_{l,t} &= \beta_{l,t-1} \left(1 + d_{l,t-1}^{-1} v_{l,t-1} \Delta T \cos \theta_{l,t-1} \right) + u_{\beta}, \end{aligned} \quad (18)$$

where $\mathbf{q}_{l,t} = [\theta_{l,t}, d_{l,t}, v_{l,t}, \beta_{l,t}]$ represents the state vector (i.e., angle, distance, velocity, and path loss) of patient l at time slot t , with $t = 1, \dots, T$. The duration of each time slot is denoted by ΔT . The term $\mathbf{u}_t = [u_{\theta}, u_d, u_v, u_{\beta}]$ represents the state process noise.

Following a similar derivation procedure, we derive the state model for up (+) and down(-) movement:

$$\begin{aligned} \theta_{l,t} &= \theta_{l,t-1} \pm d_{l,t-1}^{-1} v_{l,t-1} \Delta T \sin \left(\frac{\pi}{2} + \theta_{l,t-1} \right) + u_{\theta}, \\ d_{l,t} &= d_{l,t-1} \mp v_{l,t-1} \Delta T \cos \left(\frac{\pi}{2} + \theta_{l,t-1} \right) + u_d, \\ v_{l,t} &= v_{l,t-1} + u_v, \\ \beta_{l,t} &= \beta_{l,t-1} \left(1 + d_{l,t-1}^{-1} v_{l,t-1} \Delta T \cos \left(\frac{\pi}{2} + \theta_{l,t-1} \right) \right) + u_{\beta}. \end{aligned} \quad (19)$$

Additionally, if the patient has no movement, we have the following model:

$$\begin{aligned} \theta_{l,t} &= \theta_{l,t-1} + u_{\theta}, & d_{l,t} &= d_{l,t-1} + u_d, \\ v_{l,t} &= v_{l,t-1} + u_v, & \beta_{l,t} &= \beta_{l,t-1} + u_{\beta}. \end{aligned} \quad (20)$$

Remark 2. We assume the patient is confined to a limited area of movement, such as the dimensions of a bed or a designated region within a room. Once the patient reaches the boundary of this region, they must transition to an alternative state model that accounts for restricted mobility and prevents further physical displacement.

By denoting the measured parameters as $\mathbf{r}_{l,t} = [\hat{z}_{l,t}, \hat{d}_{l,t}, \hat{v}_{l,t}]$ and the measurement noise as $\mathbf{z}_t = [z_{\theta}, z_r, z_{\mu}]$, we can summarise the state model and the measurement model as follows:

$$\begin{cases} \text{State model:} & \mathbf{q}_{l,t} = \mathbf{g}_1(\mathbf{q}_{l,t-1}) + \mathbf{u}_t, \\ \text{Measurement model:} & \mathbf{r}_{l,t} = \mathbf{g}_2(\mathbf{q}_{l,t}) + \mathbf{z}_t, \end{cases} \quad (21)$$

where $\mathbf{g}_1(\cdot)$ is the state transition function, and $\mathbf{g}_2(\cdot)$ is the measurement function. As $\mathbf{u}_t$ and $\mathbf{z}_t$ are noise vectors with

zero-mean Gaussian distribution, their covariance matrices can be formulated by

$$\mathbf{Q}_1 = \text{diag}(\sigma_\theta^2, \sigma_d^2, \sigma_v^2, \sigma_\beta^2), \quad \mathbf{Q}_2 = \text{diag}(\sigma_1^2 \mathbf{1}, \sigma_2^2, \sigma_3^2), \quad (22)$$

where the formulas for calculating σ_1^2 , σ_2^2 and σ_3^2 are given in [54, Eq. (24)].

B. IMM Filtering for Patient Motion Tracking

While the extended Kalman filter (EKF) is effective in tracking and predicting a patient's state when their motion follows a single, predefined kinematic model, it lacks adaptability when the patient's movement involves transitions between multiple kinematic models [55], [56]. This limitation arises because the EKF operates under the assumption of a single, continuous motion model, making it suboptimal for scenarios where patients exhibit different movement patterns over time, such as rolling over in bed, sitting up, or adjusting their posture.

To address this limitation, the IMM filter is employed. The IMM filter enhances tracking accuracy by dynamically adapting to model transitions, ensuring robust state estimation even when the patient switches between different motion modes [57], [58]. The IMM filter consists of three key steps: model interaction, elementary filtering, and model probability updating. The EKF is utilised in the second step to estimate the state of each motion model, while the first and third steps facilitate the interaction between different models to achieve a more accurate overall estimation. While the detailed steps can be found in [59]–[61], the key steps are outlined below.

1) *Model Interaction*: To effectively handle uncertainty in patient motion, the state estimates and the covariance matrices from different motion models must be optimally combined. This combination is achieved through mixing weights, which determine how much effect that each model has on the overall state estimate. The mixing weights are derived from two key probabilities: the model probability $p_{t-1}^{(j)}$, which represents the likelihood that model j was correct in the previous time step, and the transition probability $\pi_{j,i}$, which defines the probability of switching from model j to model i . The mixing weights $p_{t-1}^{(ij)}$ are computed as:

$$p_{t-1}^{(ij)} = \frac{\pi_{j,i} p_{t-1}^{(j)}}{\sum_{j=1}^M \pi_{j,i} p_{t-1}^{(j)}}, \quad (23)$$

where M represents the total number of motion models.

Using these mixing weights, the predicted state and covariance matrix for model i are given by:

$$\bar{\mathbf{M}}_{t-1}^{(i)} = \sum_{j=1}^M p_{t-1}^{(ij)} \left[\hat{\mathbf{M}}_{t-1}^{(j)} + (\bar{\mathbf{q}}_{t-1}^{(i)} - \bar{\mathbf{q}}_{t-1}^{(j)}) (\bar{\mathbf{q}}_{t-1}^{(i)} - \bar{\mathbf{q}}_{t-1}^{(j)})^T \right], \quad (24)$$

where

$$\bar{\mathbf{q}}_{t-1}^{(i)} = \sum_{j=1}^M \hat{\mathbf{q}}_{t-1}^{(j)} p_{t-1}^{(ij)}. \quad (25)$$

2) *Elementary Filtering*: Once the mixed state estimates are determined, each model independently implements EKF to recursively estimate the patient's state. The EKF involves six key steps [62]:

i). State Prediction:

$$\hat{\mathbf{q}}_{t|t-1}^{(i)} = g_1^{(i)} \left(\hat{\mathbf{q}}_{t-1}^{(i)} \right), \quad (26)$$

ii). Linearisation:

$$\mathbf{G}_{1,t-1}^{(i)} = \left. \frac{\partial g_1^{(i)}}{\partial \mathbf{q}} \right|_{\mathbf{q}=\hat{\mathbf{q}}_{t-1}^{(i)}}, \quad \mathbf{G}_{2,t}^{(i)} = \left. \frac{\partial g_2^{(i)}}{\partial \mathbf{q}} \right|_{\mathbf{q}=\hat{\mathbf{q}}_{t|t-1}^{(i)}}, \quad (27)$$

iii). Covariance Matrix Prediction:

$$\mathbf{M}_{t|t-1}^{(i)} = \mathbf{G}_{1,t-1}^{(i)} \bar{\mathbf{M}}_{t-1}^{(i)} \mathbf{G}_{1,t-1}^{(i)H} + \mathbf{Q}_1, \quad (28)$$

iv). Kalman Gain Calculation:

$$\mathbf{S}_t^{(i)} = \mathbf{G}_{2,t}^{(i)} \mathbf{M}_{t|t-1}^{(i)} \mathbf{G}_{2,t}^{(i)H} + \mathbf{Q}_2, \quad (29)$$

$$\mathbf{K}_t^{(i)} = \mathbf{M}_{t|t-1}^{(i)} \mathbf{G}_{2,t}^{(i)H} \left(\mathbf{S}_t^{(i)} \right)^{-1}, \quad (30)$$

v). State Update:

$$\bar{\mathbf{r}}_t^{(i)} = \mathbf{r}_t^{(i)} - g_2^{(i)} \left(\hat{\mathbf{q}}_{t|t-1}^{(i)} \right), \quad \hat{\mathbf{q}}_t^{(i)} = \hat{\mathbf{q}}_{t|t-1}^{(i)} + \mathbf{K}_t^{(i)} \bar{\mathbf{r}}_t^{(i)}, \quad (31)$$

vi). Covariance Matrix Update:

$$\hat{\mathbf{M}}_t^{(i)} = \left(\mathbf{I} - \mathbf{K}_t^{(i)} \mathbf{G}_{2,t}^{(i)} \right) \mathbf{M}_{t|t-1}^{(i)}, \quad (32)$$

3) *Model Probability Update*: The final step of the IMM filter is updating the probability of each model given the most recent measurement. This is computed using the likelihood function:

$$\Lambda_t^{(i)} = \frac{\exp \left(-\frac{1}{2} \bar{\mathbf{r}}_t^{(i)T} \left(\mathbf{S}_t^{(i)} \right)^{-1} \bar{\mathbf{r}}_t^{(i)} \right)}{\sqrt{|2\pi \mathbf{S}_t^{(i)}|}}, \quad (33)$$

$$p_t^{(i)} = \frac{\Lambda_t^{(i)} \sum_{j=1}^M \pi_{j,i} p_{t-1}^{(j)}}{\sum_{i=1}^M \left(\Lambda_t^{(i)} \sum_{j=1}^M \pi_{j,i} p_{t-1}^{(j)} \right)}. \quad (34)$$

As such, the predicted state is given by:

$$\hat{\mathbf{q}}_t = \sum_{i=1}^M p_t^{(i)} \hat{\mathbf{q}}_t^{(i)}, \quad (35)$$

and therefore the MSE matrix can be obtained via

$$\hat{\mathbf{M}}_t = \sum_{i=1}^M p_t^{(i)} \left[\hat{\mathbf{M}}_t^{(i)} + \left(\hat{\mathbf{q}}_t^{(i)} - \hat{\mathbf{q}}_t \right) \left(\hat{\mathbf{q}}_t^{(i)} - \hat{\mathbf{q}}_t \right)^T \right]. \quad (36)$$

IV. PERFORMANCE INDICATORS

This section outlines the key performance indicators used to evaluate the system's performance.

A. Semantic Communication

The semantic transmission rate is defined as the quantity of bits received by the medical device following the extraction of semantic information, and thus, the formulation is expressed as [63]:

$$S_k = \frac{\iota}{\rho_k} \log_2 (1 + \gamma_k), \quad (37)$$

where the parameter $\rho_k = \frac{\text{len}(c_k)}{\text{len}(m_k)}$, $0 \leq \rho_k \leq 1$ represents the semantic extraction ratio, and $\text{len}(\cdot)$ denotes the length of the message. Here, ι is a scalar value converting the word-to-bit ratio. It is important to note that semantic communication in the proposed framework is optional. When $\rho_k = 1$, semantic communication is disabled, and the system defaults to conventional transmission. For example, in critical-care scenarios, this ensures that raw vital sign data (i.e., the original messages) are transmitted continuously without semantic compression, thereby guaranteeing maximum accuracy and reliability. In addition, the lower bound of ρ_k , denoted by $\rho_{k,LB}$, has been derived in [64], [65]:

$$\rho_{k,LB} = \frac{1}{1 - \ln Q + \sum_g^G w_{g,k} \log p_{g,k}}, \quad (38)$$

where Q represents the global lower bound of all the individual Bilingual Evaluation Understudy (BLEU) scores. The BLEU score evaluates how closely the reconstructed message, obtained from the received semantic information, matches the original message. Additionally, $w_{g,k}$ denotes the weight assigned to the g -grams, where G is the total number of g -grams required to represent a sentence. The precision score $p_{g,k}$ quantifies the accuracy of the message recovered by medical device k .

To prevent information leakage, we evaluate the worst-case semantic secrecy rate (SSR). A larger SSR means a lower chance of information leakage. In the worst-case scenario, where the patient has an extensive KB similar to that of the robot and the medical devices, i.e., $\rho_{l|k} = \rho_k$, the semantic transmission rate for the l -th patient related to the k -th medical device is determined as follows:

$$S_{l|k} = \frac{\iota}{\rho_k} \log_2 (1 + \Gamma_{l|k}). \quad (39)$$

In this way, the worst-case SSR of the k -th medical device is formulated by

$$SSR_k = \max \left(\min_{l \in L} [S_k - S_{l|k}], 0 \right). \quad (40)$$

B. Computing and Power Consumption

Semantic information extraction from conventional messages relies heavily on advanced machine learning models, such as Transformers, which are computationally intensive. Therefore, the power consumption of computing plays a pivotal role in semantic communication and must be explicitly accounted for in the overall transmission power budget. In [25], [63], the authors established and validated a relationship between the semantic extraction ratio ρ_k and the corresponding computing power, showing a positive correlation within a specific range of ρ_k . Accordingly, the computing power function

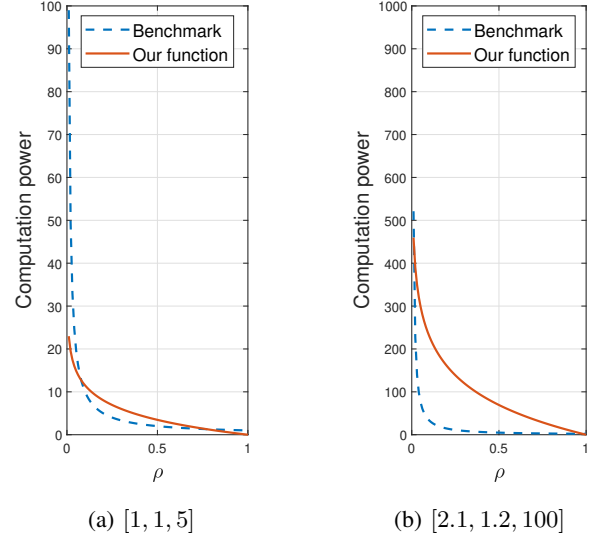


Figure 4: An illustration of the approximation of [63, Eq. (10)] using (41). The parameters are given as $[C_1, C_2, F]$.

was modelled in a piecewise manner. However, identifying the appropriate range and slope of each segment necessitates extensive hyperparameter tuning. To avoid manual hyperparameter selection, we instead adopt a natural logarithmic function to model the computing power:

$$P_{\text{comp}} = \sum_{k=1}^K -F \ln(\rho_k), \quad (41)$$

where F is an energy-efficiency coefficient that is dependent on the CPU design. To demonstrate the difference between our function and [63, Eq. (10)], we assume $K = 1$ and re-state [63, Eq. (10)] as:

$$P_{\text{comp}} = C_1 \rho_k^{-C_2},$$

where C_1 and C_2 are positive constants. By setting C_1, C_2, F to their corresponding values, we obtain Figs. 4.

The preference for the proposed function is motivated by two key reasons:

- i). Eq. (41) provides a smoother increase of power when ρ decreases from a large value. Whereas [63, Eq. (10)] only experiences large increments of power when ρ is relatively small.
- ii). Eq. (41) involves only a single hyperparameter, whereas [63, Eq. (10)] requires two. Moreover, when $\rho = 1$, indicating that no semantic extraction is performed, Eq. (41) yields zero computing power, while [63, Eq. (10)] produces a small positive value. Since no semantic processing occurs when $\rho = 1$, the computing power should ideally be zero.

On the other hand, the communication and sensing power consumption at the transmitter side is given by [66]:

$$P_{\text{c\&s}} = \text{Tr} \left(\sum_{k=1}^K \mathbf{W}_k + \sum_{l=1}^L \mathbf{R}_l \right). \quad (42)$$

Additionally, the overall transmission power consumption is limited to the power budget:

$$P_{\text{comp}} + P_{\text{c\&s}} \leq P_t, \quad (43)$$

with P_t being the total power budget.

C. Sensing

To accurately track the motion of the patient and estimate vital sign information from the echo signal, it is essential to precisely direct the beams toward the patient's chest, corresponding to the optimal measurement model in (21). This necessitates assessing angle estimation accuracy by calculating and optimising the MSE between the detected and true angles, thereby ensuring that the beams are effectively aligned with the correct directions. However, deriving a closed-form expression for MSE can be challenging. As an alternative, we utilise the Cramér-Rao bound (CRB) to assess the performance of patient sensing in a static scenario. For unbiased estimators, the CRB provides a theoretical lower bound on the MSE, i.e., $\text{MSE}(\hat{\theta}) \geq \text{CRB}(\theta)$, offering an analytically tractable means of performance evaluation. Since the CRB formulation has been derived in [67], we provide a brief description below.

Defining the parameters to be estimated as $\xi_l = [\theta_l, \beta_l]$, the Fisher information matrix (FIM) about ξ_l is given by:

$$\mathbf{J}_l = \begin{bmatrix} J_{\theta_l \theta_l} & \mathbf{J}_{\theta_l \beta_l} \\ \mathbf{J}_{\theta_l \beta_l}^T & J_{\beta_l \beta_l} \end{bmatrix}. \quad (44)$$

Let us define: $\mathbf{B}_l = \mathbf{a}(\theta_l) \mathbf{a}^H(\theta_l)$ and $\dot{\mathbf{B}}_{\theta_l} = \frac{\partial \mathbf{B}_l}{\partial \theta_l}$. Hence, we can obtain the following equations:

$$J_{\theta_l \theta_l} = \frac{2T|\beta_l|^2}{\sigma_e^2} \text{Tr} \left(\dot{\mathbf{B}}_{\theta_l} \mathbf{R}_x \dot{\mathbf{B}}_{\theta_l}^H \right), \quad (45)$$

$$\mathbf{J}_{\beta_l \beta_l} = \frac{2T}{\sigma_e^2} \text{Tr} \left(\mathbf{B}_l \mathbf{R}_x \mathbf{B}_l^H \right) \mathbf{I}, \quad (46)$$

$$\mathbf{J}_{\theta_l \beta_l} = \frac{2T\beta_l^*}{\sigma_e^2} \Re \left\{ \text{Tr} \left(\mathbf{B}_l \mathbf{R}_x \dot{\mathbf{B}}_{\theta_l}^H \right) \right\} [1, j]. \quad (47)$$

Therefore, the CRB of θ_l is formulated by

$$\text{CRB}(\theta_l) = \mathbf{J}_l^{-1} [1, 1] = \left(J_{\theta_l \theta_l} - \mathbf{J}_{\theta_l \beta_l} \mathbf{J}_{\beta_l \beta_l}^{-1} \mathbf{J}_{\theta_l \beta_l}^T \right)^{-1}, \quad (48)$$

where $\mathbf{A}_{[1,1]}$ means the element of matrix $\mathbf{A}$ in row 1 and column 1.

Remark 3. The advantage of semantic communication over traditional communication methods, while maintaining the sensing performance, will be demonstrated in Section VI.A.

For tracking performance, we focus on the root mean square error (RMSE) of angle and distance, which are given by:

$$\text{RMSE}_{\theta,l} = \sqrt{(\theta_l - \hat{\theta}_l)^2}, \quad \text{RMSE}_{d,l} = \sqrt{(d_l - \hat{d}_l)^2}, \quad (49)$$

Finally, to evaluate the accuracy of vital sign estimation, we adopt the RMSE as a standard performance metric. Specifically, the RMSEs for respiration and heartbeat frequency estimates are defined as

$$\text{RMSE}_{r,l} = \sqrt{(f_{r,l} - \hat{f}_{r,l})^2}, \quad \text{RMSE}_{h,l} = \sqrt{(f_{h,l} - \hat{f}_{h,l})^2}, \quad (50)$$

Having defined these key performance indicators for semantic communication, power consumption, and sensing accuracy, we are now equipped to formulate an optimisation problem that balances these competing objectives.

V. PROBLEM FORMULATION AND ALGORITHM DESIGN

This section formulates the optimisation problem and proposes the corresponding algorithms.

A. Problem Formulation

Since the medical devices are movable within a limited range (e.g., to facilitate better treatment), the communication channels are inherently uncertain. Consequently, the objective is to maximise the worst-case SSR across all medical devices to guarantee robustness. Additionally, to ensure precise beam alignment with the patients' chest angles, we minimise the sum CRB of the patient angles, as a lower CRB value corresponds to reduced sensing error. This dual optimisation enhances semantic secrecy while improving the precision of patient angle estimates.

The optimisation problem is subject to several constraints. Firstly, the semantic extraction ratio ρ_k is lower-bounded by $\rho_{k,LB}$ as shown in (38). Furthermore, the constraint of the transmit power budget must be satisfied, as defined in (43). Additionally, the positive semi-definite property of matrices must be satisfied. Moreover, to facilitate signal transmission through a single-stream transmit beamforming for each medical device, a rank-one constraint is imposed. It is worth noting that the inclusion of this constraint eliminates the need for more complex transceiver schemes, as discussed in [68]. Hence, the optimisation problem is formulated as:

$$\max_{\mathbf{W}_k, \mathbf{R}_l, \rho_k} \kappa_1 \min_{k \in K} (\text{SSR}_k) - \kappa_2 \sum_{l=1}^L \text{CRB}(\theta_l) \quad (51a)$$

$$\text{s.t.} \quad \rho_{k,LB} \leq \rho_k \leq 1, \forall k, \quad (51b)$$

$$P_{\text{comp}} + P_{\text{c\&s}} \leq P_t, \quad (51c)$$

$$\mathbf{W}_k \succeq 0, \mathbf{W}_k = \mathbf{W}_k^H, \forall k, \quad (51d)$$

$$\mathbf{R}_l \succeq 0, \mathbf{R}_l = \mathbf{R}_l^H, \forall l, \quad (51e)$$

$$\text{rank}(\mathbf{W}_k) = 1, \forall k, \quad (51f)$$

where κ_1 and κ_2 are the weights.

B. Algorithm Design

To simplify the objective function, we first transfer (51) into the following form:

$$\max_{\mathbf{W}_k, \mathbf{R}_l, \rho_k, \lambda} \kappa_1 \lambda - \kappa_2 \sum_{l=1}^L \text{CRB}(\theta_l) \quad (52a)$$

$$\text{s.t.} \quad S_k - S_{l|k} \geq \lambda, \forall k, \forall l, \quad (52b)$$

$$(51b) - (51f). \quad (52c)$$

Since both medical devices and patients move within a limited range, we assume that the channel uncertainty is confined within a bounded spherical region [69]–[71]:

$$\mathcal{H}_i := \left\{ \left(\hat{\mathbf{h}}_i + \mathbf{u}_i \right)^H \mid \|\mathbf{u}_i\| \leq \varepsilon_i \right\}, \quad i \in [k, l], \quad (53)$$

where $\varepsilon_i \geq 0$ corresponds to the radius of $\mathcal{H}_i$. The estimated channel is denoted by $\hat{\mathbf{h}}_i$, while $\mathbf{u}_i$ represents the error vector limited by ε_i . Next, we deal with the non-concave constraint (52b). The upper and lower bounds of $|\mathbf{h}_i \sum_{k=1}^K \mathbf{w}_k|^2$ are given by

$$\begin{aligned} \left| \mathbf{h}_i \sum_{k=1}^K \mathbf{w}_k \right|^2 &\leq \hat{\mathbf{h}}_i \sum_{k=1}^K \mathbf{w}_k \hat{\mathbf{h}}_i^H + 2\varepsilon_i \left\| \sum_{k=1}^K \mathbf{w}_k \hat{\mathbf{h}}_i^H \right\|, \\ \left| \mathbf{h}_i \sum_{k=1}^K \mathbf{w}_k \right|^2 &\geq \hat{\mathbf{h}}_i \sum_{k=1}^K \mathbf{w}_k \hat{\mathbf{h}}_i^H - 2\varepsilon_i \left\| \sum_{k=1}^K \mathbf{w}_k \hat{\mathbf{h}}_i^H \right\|. \end{aligned} \quad (54)$$

Proof. By neglecting the term $\mathbf{u}_i \sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \mathbf{u}_i^H$, we derive the upper bound via

$$\begin{aligned} \left| \mathbf{h}_i \sum_{k=1}^K \mathbf{w}_k \right|^2 &= \left(\hat{\mathbf{h}}_i + \mathbf{u}_i \right) \sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \left(\hat{\mathbf{h}}_i + \mathbf{u}_i \right)^H \\ &\stackrel{(a)}{\approx} \hat{\mathbf{h}}_i \sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \hat{\mathbf{h}}_i^H + 2\Re \left\{ \mathbf{u}_i \sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \hat{\mathbf{h}}_i^H \right\} \\ &\stackrel{(b)}{\leq} \hat{\mathbf{h}}_i \sum_{k=1}^K \mathbf{w}_k \hat{\mathbf{h}}_i^H + 2 \left\| \mathbf{u}_i \sum_{k=1}^K \mathbf{w}_k \hat{\mathbf{h}}_i^H \right\| \\ &\stackrel{(c)}{\leq} \hat{\mathbf{h}}_i \sum_{k=1}^K \mathbf{w}_k \hat{\mathbf{h}}_i^H + 2\varepsilon_i \left\| \sum_{k=1}^K \mathbf{w}_k \hat{\mathbf{h}}_i^H \right\|, \end{aligned}$$

where step (a) follows from expanding the quadratic form and discarding the error term $\mathbf{u}_i \mathbf{W}_k \mathbf{u}_i^H$. Step (b) uses the inequality $\Re\{z\} \leq |z|$ for any complex number z , and step (c) follows from the Cauchy–Schwarz inequality. The proof of the corresponding lower bound follows similarly by using $\Re\{z\} \geq -|z|$. $\square$

In a similar way, the lower and upper bounds of $|\mathbf{h}_i \sum \mathbf{r}_l|^2$ can be obtained. As such, the lower bound of (52b) is found and shown in (55). However, (55) remains non-convex, primarily due to the terms $\max - \log_2(\text{Term } 2)$ and $\max - \log_2(\text{Term } 3)$. Inspired by [72], to address this issue, we consider the following change of variables in (55):

$$\begin{aligned} e^{a_k} &\triangleq \text{Term } 1, & e^{b_k} &\triangleq \text{Term } 2, \\ e^{c_k} &\triangleq \text{Term } 3, & e^{d_k} &\triangleq \text{Term } 4. \end{aligned} \quad (56)$$

By introducing exponential terms, the expression such as $\max - \log_2(\text{Term } 2)$ is effectively transformed into $\max - b_k \log_2(e)$, which is a linear and convex function. Thus, (55) can be written in the form of:

$$\frac{\ell}{\rho_k} \log_2 \left(\frac{e^{a_k}}{e^{b_k}} \right) - \frac{\ell}{\rho_k} \log_2 \left(\frac{e^{c_{l|k}}}{e^{d_{l|k}}} \right) \geq \lambda. \quad (57)$$

Finally, to tackle the non-concave function CRB in the objective function, we transform it into a constraint of

$$\begin{bmatrix} J_{\theta_l \theta_l} - U_l & \mathbf{J}_{\theta_l \beta_l} \\ \mathbf{J}_{\theta_l \beta_l}^T & \mathbf{J}_{\beta_l \beta_l} \end{bmatrix} \succeq 0, \forall l, \quad (58)$$

with the CRB part in the objective function becomes $-\kappa_2 \left(\sum_{l=1}^L U_l^{-1} \right)$ and $U_l \geq 0$ is a new variable introduced.

Hence, (52) can be reformulated in the following way:

$$\max_{\Psi} \kappa_1 \lambda - \kappa_2 \left(\sum_{l=1}^L U_l^{-1} \right) \quad (59a)$$

$$\text{s.t. } \frac{\ell}{\rho_k} \log_2 e^{a_k - b_k + d_{l|k} - c_{l|k}} \geq \lambda, \forall k, \forall l, \quad (59b)$$

$$\text{Term } 1 \geq e^{a_k}, \forall k, \quad (59c)$$

$$\text{Term } 2 \leq e^{b_k}, \forall k, \quad (59d)$$

$$\text{Term } 3 \leq e^{c_{l|k}}, \forall l, \forall k, \quad (59e)$$

$$\text{Term } 4 \geq e^{d_{l|k}}, \forall k, \forall l, \quad (59f)$$

$$(51b), (51c), (51d), (51e), (51f), (58), \quad (59g)$$

where $\Psi = [\mathbf{W}_k, \mathbf{R}_l, U_l, \rho_k, \lambda, a_k, b_k, c_{l|k}, d_{l|k}]$. In addition, it has been verified in [72] that all the inequalities in (59c) to (59f) hold with equalities at the optimal points.

Nevertheless, constraints (59d) and (59e) are still non-convex. We consider using the first-order Taylor Expansion. Let us introduce two more equations:

$$\begin{aligned} \hat{b}_k &= \ln \left(\hat{\mathbf{h}}_k \sum_{k'=1, k' \neq k}^K \mathbf{w}_{k'}^i \hat{\mathbf{h}}_k^H + 2\varepsilon_k \left\| \sum_{k'=1, k' \neq k}^K \mathbf{w}_{k'}^i \hat{\mathbf{h}}_k^H \right\| \right. \\ &\quad \left. + \hat{\mathbf{h}}_k \sum_{l=1}^L \mathbf{r}_l^i \hat{\mathbf{h}}_k^H + 2\varepsilon_k \left\| \sum_{l=1}^L \mathbf{r}_l^i \hat{\mathbf{h}}_k^H \right\| + \sigma_c^2 \right), \end{aligned} \quad (60)$$

$$\begin{aligned} \hat{c}_{l|k} &= \ln \left(\hat{\mathbf{h}}_l \sum_{k=1}^K \mathbf{w}_k^i \hat{\mathbf{h}}_l^H + 2\varepsilon_l \left\| \sum_{k=1}^K \mathbf{w}_k^i \hat{\mathbf{h}}_l^H \right\| \right. \\ &\quad \left. + \hat{\mathbf{h}}_l \sum_{l=1}^L \mathbf{r}_l^i \hat{\mathbf{h}}_l^H + 2\varepsilon_l \left\| \sum_{l=1}^L \mathbf{r}_l^i \hat{\mathbf{h}}_l^H \right\| + \sigma_r^2 \right). \end{aligned} \quad (61)$$

By applying (60) and (61), (59d) and (59e) can be replaced by the convex functions (62) and (63), which are shown on the top of the next page. To solve the optimisation problem, we drop the rank-one constraint, then split the problem into three sub-problems and use the alternating optimising method:

1) *Sub-problem 1:* With given ρ_k in problem (59), the joint sensing and communication beamforming optimisation problem can be given by

$$\max_{\Psi} \kappa_1 \lambda - \kappa_2 \left(\sum_{l=1}^L U_l^{-1} \right) \quad (64a)$$

$$\text{s.t. } (51d), (51e), (58), (59b), (59c), (59f), (62), (63), \quad (64b)$$

where $\Psi = [\mathbf{W}_k, \mathbf{R}_l, U_l, \lambda, a_k, b_k, c_{l|k}, d_{l|k}]$. Optimisation problem (64) is concave and can be effectively solved via the standard convex optimisation tool, such as CVX [73].

2) *Sub-problem 2:* When semantic communication is selected (i.e., $\rho_k \neq 1$), given $\Psi = [\mathbf{W}_k, \mathbf{R}_l, U_l, \lambda, a_k, b_k, c_{l|k}, d_{l|k}]$, as well as denoting $D_k = \log_2 e^{a_k - b_k + d_{l|k} - c_{l|k}}$, the semantic extraction ratio optimisation problem can be given by

$$\max_{\rho_k} \frac{\ell}{\rho_k} D_k \quad (65a)$$

$$\text{s.t. } (51b), (51c), (59b), \quad (65b)$$

$$\begin{aligned}
& \frac{\iota}{\rho_k} \log_2 \left(\underbrace{\frac{\hat{\mathbf{h}}_k \sum_{k=1}^K \mathbf{W}_k \hat{\mathbf{h}}_k^H - 2\varepsilon_k \left\| \sum_{k=1}^K \mathbf{W}_k \hat{\mathbf{h}}_k^H \right\| + \hat{\mathbf{h}}_k \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_k^H - 2\varepsilon_k \left\| \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_k^H \right\| + \sigma_c^2}{\hat{\mathbf{h}}_k \sum_{k'=1, k' \neq k}^K \mathbf{W}_{k'} \hat{\mathbf{h}}_k^H + 2\varepsilon_k \left\| \sum_{k'=1, k' \neq k}^K \mathbf{W}_{k'} \hat{\mathbf{h}}_k^H \right\| + \hat{\mathbf{h}}_k \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_k^H + 2\varepsilon_k \left\| \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_k^H \right\| + \sigma_c^2}}_{\text{Term 1}} \right) \\
& - \frac{\iota}{\rho_k} \log_2 \left(\underbrace{\frac{\hat{\mathbf{h}}_l \sum_{k=1}^K \mathbf{W}_k \hat{\mathbf{h}}_l^H + 2\varepsilon_l \left\| \sum_{k=1}^K \mathbf{W}_k \hat{\mathbf{h}}_l^H \right\| + \hat{\mathbf{h}}_l \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_l^H + 2\varepsilon_l \left\| \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_l^H \right\| + \sigma_r^2}{\hat{\mathbf{h}}_l \sum_{k'=1, k' \neq k}^K \mathbf{W}_{k'} \hat{\mathbf{h}}_l^H - 2\varepsilon_l \left\| \sum_{k'=1, k' \neq k}^K \mathbf{W}_{k'} \hat{\mathbf{h}}_l^H \right\| + \hat{\mathbf{h}}_l \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_l^H - 2\varepsilon_l \left\| \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_l^H \right\| + \sigma_r^2}}_{\text{Term 2}} \right) \\
& \geq \lambda. \tag{55}
\end{aligned}$$

$$\hat{\mathbf{h}}_k \sum_{k'=1, k' \neq k}^K \mathbf{W}_{k'} \hat{\mathbf{h}}_k^H + 2\varepsilon_k \left\| \sum_{k'=1, k' \neq k}^K \mathbf{W}_{k'} \hat{\mathbf{h}}_k^H \right\| + \hat{\mathbf{h}}_k \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_k^H + 2\varepsilon_k \left\| \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_k^H \right\| + \sigma_c^2 \leq e^{\hat{b}_k} (b_k - \hat{b}_k + 1), \tag{62}$$

$$\hat{\mathbf{h}}_l \sum_{k=1}^K \mathbf{W}_k \hat{\mathbf{h}}_l^H + 2\varepsilon_l \left\| \sum_{k=1}^K \mathbf{W}_k \hat{\mathbf{h}}_l^H \right\| + \hat{\mathbf{h}}_l \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_l^H + 2\varepsilon_l \left\| \sum_{l=1}^L \mathbf{R}_l \hat{\mathbf{h}}_l^H \right\| + \sigma_r^2 \leq e^{\hat{c}_{l|k}} (c_{l|k} - \hat{c}_{l|k} + 1). \tag{63}$$

Proposition 1. The optimal value of ρ_k , denoted by ρ_k^* , is equal to $\min \left(\max \left(\frac{\iota D_k}{\eta^* F}, \rho_{LB,k} \right), \rho_{UB,k} \right)$, where η is the Lagrange multiplier and η^* is the optimal value.

Proof. See Appendix A $\square$

By applying **Proposition 1**, the optimal value of ρ_k is found.

3) *Sub-problem 3:* As a last step, we use Gaussian randomisation to recover rank-one solutions.

The detailed steps for solving problem (59) are outlined in Algorithm 1. The computational complexity of Algorithm 1 is $\mathcal{O}(I_1 I_2 (KN^2)^2)$, where I_1 represents the number of iterations in the outer loop, and I_2 denotes the number of iterations required to solve the problem (64). The term $(KN^2)^2$ arises because the problem involves K matrices of size $N \times N$, and the quadratic nature of the optimisation problem results in the squared term.

VI. NUMERICAL RESULTS

In this section, we present numerical results to evaluate the performance of the proposed framework. The robot is equipped with a ULA consisting of 8 antennas with half-wavelength spacing. The patients are positioned at $(-35^\circ, 2\text{m})$, $(5^\circ, 5\text{m})$, and $(25^\circ, 1\text{m})$, while the medical devices are located at $(-30^\circ, 8\text{m})$ and $(20^\circ, 4\text{m})$. The noise power is fixed at -60 dBm, and the total transmit power is constrained to 20 dBm. Each channel $\mathbf{h}_i$, $i \in [k, l]$ consists of both line-of-sight (LoS) and NLoS components. In (37), the parameter ι is set to 1.1, and the trade-off coefficient κ is set to 0.5. The lower bounds of the semantic extraction ratios ρ_1 , ρ_2 , and ρ_3 are randomly selected based on the dataset¹, which has an average value of 0.084 with

Algorithm 1 IMM Filter Procedure and Iterative Sensing, Communication, and Semantic Optimisation Algorithm

- 1: **repeat**
- 2: Perform the model interaction step from the IMM Filter.
- 3: Execute the state prediction, linearization, and MSE prediction steps from the elementary filtering process.
- 4: Compute the predicted state estimate for predictive beamforming: $\hat{\mathbf{q}}_{t|t-1} = \sum_{i=1}^M \hat{\mathbf{q}}_{t|t-1}^{(i)} p_{t-1}^{(i)}$.
- 5: Initialize variables: ρ_k , $\hat{c}_{l|k,i}$, and $\hat{b}_{k,i}$.
- 6: **repeat**
- 7: **repeat**
- 8: Solve the optimization problem (64) to update $\hat{c}_{l|k,i+1}$ and $\hat{b}_{k,i+1}$.
- 9: **until** $|\hat{c}_{l|k,i+1} - \hat{c}_{l|k,i}| \leq \varrho_1$ **and** $|\hat{b}_{k,i+1} - \hat{b}_{k,i}| \leq \varrho_2$.
- 10: **repeat**
- 11: Solve (65) using **Proposition 1** and update ρ_k .
- 12: **until**
- 13: **until** $|\mathbf{W}^{i+1} - \mathbf{W}^i| \leq \varrho_3$ **and** $|\mathbf{R}^{i+1} - \mathbf{R}^i| \leq \varrho_4$.
- 14: Apply Gaussian randomisation.
- 15: Execute the Kalman gain calculation, state update, and MSE matrix update steps from the elementary filtering process.
- 16: Perform the model probability update step.
- 17: **until**

standard deviation of 0.089. Furthermore, in (53), the channel uncertainty bound is specified as $\varepsilon_i = 0.01$. The length of the time slot ΔT is set to 0.1 seconds. For the vital sign modelling, the respiration and heartbeat frequencies are generated in the range $0.1 - 0.5$ Hz and $1 - 2.5$ Hz, respectively, and amplitudes are generated around $A_r = 5$ mm and $A_h = 1$

¹ Available at <https://www.kaggle.com/datasets/yangtony1999/medical-dataset-for-amplitude-and-frequency-analysis>.

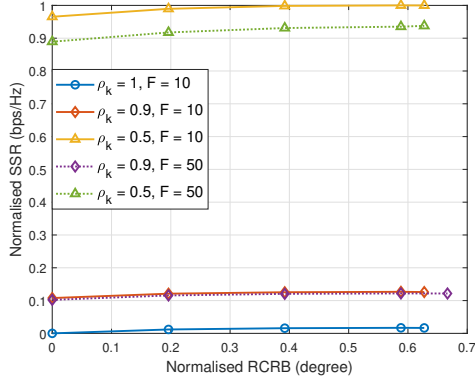


Figure 5: Normalised SSR versus normalised RCRB for different values of ρ_k and F .

mm, respectively. The patients may be moving at a speed of 0.1 m/s. The state transition noise covariance matrix is defined as $\mathbf{Q}_1 = \text{diag}(10^{-1}, 1, 10^{-2}, 10^{-4})$.

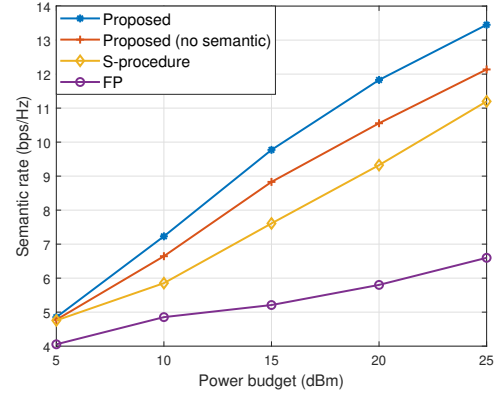
A. Discussion of Remark 3

For analytical traceability, this subsection considers a simplified scenario involving a single patient and a single medical device. To facilitate a performance comparison between traditional and semantic communication schemes, the parameter ρ_k is set to 1 to represent traditional communication, while values of 0.9 and 0.5 are employed to characterise semantic communication. The value of F is set to 10 and 50, illustrating the system performance under varying device computing capabilities. Consequently, sub-problem (65) is omitted from this analysis.

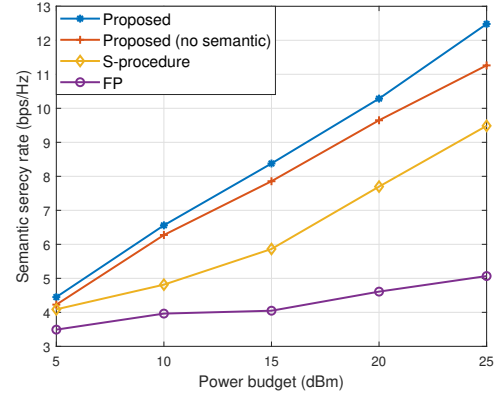
Fig. 5 illustrates the normalised SSR and the normalised root CRB (RCRB) with varying values of ρ_k and F . The plotted trends demonstrate that, for identical RCRB values, semantic communication yields improved performance over the traditional method, highlighting its ability to enhance communication efficiency without compromising sensing accuracy. Specifically, a lower ρ_k yields a more improved communication performance. Furthermore, as F increases from 10 to 50, a slight degradation in communication performance is observed. This is attributed to the increased power consumption associated with semantic feature extraction at higher F , which reduces the available power for communication and sensing, thereby slightly compromising the communication rate to maintain the sensing quality. These findings confirm the advantage of semantic communication in achieving better communication performance while maintaining sensing performance.

B. Semantic Communication Performance

In this subsection, we compare the performance of the proposed design with the methods presented in [74] and [75], respectively. The approach in [74] employs the S-procedure for robust beamforming, while [75] utilises fractional programming (FP) for non-robust performance optimisation.



(a) Semantic rate



(b) Semantic secrecy rate

Figure 6: Semantic rate and semantic secrecy rate against power budget.

As shown in Fig. 6(a), when the power budget is limited to 5 dBm, the proposed design achieves identical communication performance as the S-procedure-based design. This is primarily because the limited power budget restricts the system's ability to effectively extract semantic features. The non-robust FP method yields the lowest transmission rate due to its sensitivity to channel uncertainties. As the power budget increases, the advantages of both semantic communication and the proposed algorithm become more evident. For instance, at 20 dBm, the proposed design without semantic achieves a transmission rate approximately 10% higher than the S-procedure approach, and nearly double that of the FP method. When semantic communication is enabled, the performance gain becomes even more significant, with a 20% improvement over the S-procedure design. Notably, the proposed design maintains consistently high performance as the power budget increases, whereas the FP method shows limited improvement under higher power levels. These findings confirm the advantage of semantic communication in achieving efficient information transmission while maintaining sensing performance.

Fig. 6(b) depicts the SSR as a function of the power budget. The proposed scheme without semantic communication consistently outperforms the S-procedure and FP-based designs, with the performance gap widening as the power

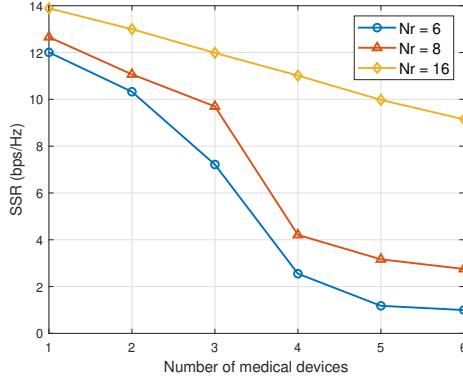


Figure 7: Semantic secrecy rate against the number of medical devices.

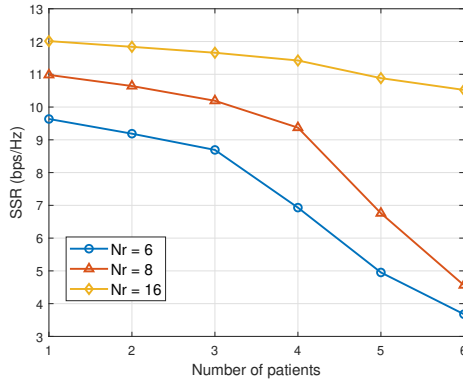


Figure 8: Semantic secrecy rate against the number of patients.

budget increases. Incorporating semantic communication further enhances performance, owing to the effective integration of semantic techniques that improve both the semantic rate and the SSR. Notably, the advantage of the proposed design becomes more pronounced at higher power levels.

Fig. 7 illustrates the average SSR under varying numbers of medical devices and different numbers of antennas. To ensure reliable communication and sensing performance, a common design constraint is that the total number of objects should remain significantly smaller than the number of antennas, i.e., $K + L \ll N$. With the number of patients fixed at $L = 3$, we vary the number of medical devices K . When $N = 6$, a sharp decline in SSR is observed from $K = 2$, indicating a violation of the $K + L \ll N$ condition. Increasing the number of antennas to $N = 8$ slightly improves performance, although a noticeable drop still occurs from $K = 3$. With $N = 16$, the SSR exhibits a gradual decline as K increases. This trend arises because adding more devices reduces the per-device resource availability, thereby slightly degrading the SSR while still maintaining performance at high levels.

Fig. 8 illustrates the average SSR for varying numbers of patients and antennas. As the number of patients increases, the SSR decreases due to the elevated risk of eavesdropping. When the condition $K + L \ll N$ is satisfied, as in the case of $N = 16$, the SSR remains relatively high. In contrast, when this

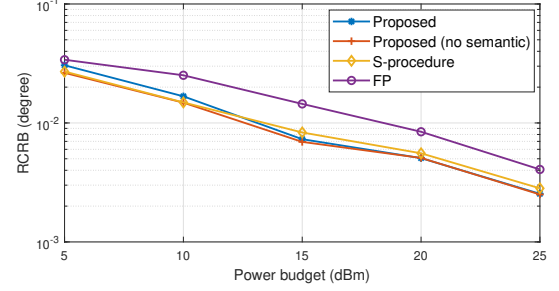


Figure 9: Sum RCRB against power budget.

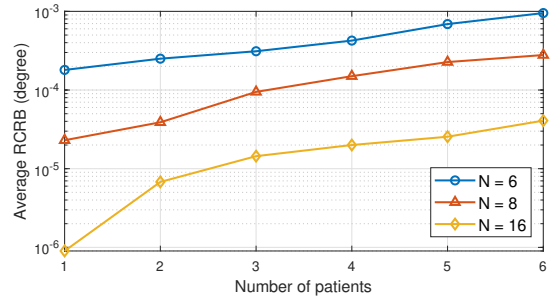


Figure 10: Average sum RCRB against number of patients.

condition is not met, as with $N = 6$, the SSR experiences a pronounced degradation. Consequently, Figs. 7 and 8 indicate that equipping the robot with a sufficiently large number of antennas is essential to maintain communication and security performance; alternatively, a cooperative robot system could be deployed to achieve the same objective.

C. Sensing and Tracking Performance

Regarding sensing performance, Fig. 9 depicts the sum RCRB values. At lower power levels, the proposed design using semantic communication yields slightly higher RCRB values than the non-semantic scheme, indicating a minor degradation in sensing accuracy due to resource sharing with semantic extraction. However, as the power budget increases, this performance gap gradually diminishes. Notably, at 15 dBm, the proposed design with and without semantic communication achieves nearly equivalent sensing performance and outperforms the S-procedure method. This behaviour confirms that the proposed framework can sustain competitive sensing accuracy under moderate-to-high power budgets while simultaneously delivering enhanced communication performance, as shown in Fig. 6.

Fig. 10 illustrates the average sum RCRB values as the number of patients increases. As expected, the sensing accuracy generally deteriorates with an increasing number of patients, primarily due to the resources being divided among more targets. This resource constraint limits the achievable resolution and detection capability for each individual patient. To address this challenge and improve sensing performance, two main approaches can be considered. First, the robot

can be equipped with a larger antenna array. For example, increasing the number of antennas from 6 to 16 reduces the RCRB from 10^{-3} to 10^{-6} , thereby enhancing sensing accuracy. Alternatively, deploying a network of cooperating robots enables distributed sensing, where multiple robots jointly process sensing data to exploit spatial diversity and cooperative gains. This observation aligns with the conclusion in Figs. 7 and 8. Such a cooperative design can effectively mitigate resource constraints, resulting in more robust communication performance and improved sensing accuracy in multi-patient and multi-device scenarios.

Table II summarises the average tracking performance over time, quantified by RMSEs shown in (49), under varying noise levels, patient speeds, and movement probabilities. As an example, a movement probability of 0.15 indicates that each of the four possible movement directions occurs with probability 0.15, while the probability of no movement is 0.4. When the noise variance $\sigma_e^2 = 1$, corresponding to relatively high noise conditions, the RMSE values are on the order of 10^{-1} . As the movement probability increases, $\text{RMSE}_{\theta,l}$ exhibits only a marginal rise, demonstrating the effective design of the IMM filter and the beamforming matrices. While $\text{RMSE}_{d,l}$ increases slightly more than $\text{RMSE}_{\theta,l}$ as patient speed increases, it remains within acceptable limits. Notably, increasing the patient's speed does not lead to higher RMSEs, which highlights the robustness of the IMM filter. Under lower noise conditions, the RMSEs decrease to the order of 10^{-2} . A similar conclusion can be found, that is, RMSEs remain stable with increasing speed but increase slightly as the movement probability rises. This slight increase in RMSEs with movement probability can be attributed to the challenge of accurately tracking more frequent movements, especially rapid ones. Nevertheless, the IMM filter maintains robust tracking performance under these dynamic conditions.

D. Vital Sign Estimation Performance

Fig. 11 presents the experimental results of multi-patient vital sign sensing. Specifically, Fig. 11(a) illustrates the time-domain respiration signals, while Fig. 11(b) depicts the time-domain heartbeat signals. These signals are obtained after applying bandpass filtering as described in (13) and (14), respectively. The corresponding ground truth values are provided in (10). As shown, the filtered signals retain the periodic characteristics of the original waveforms, thereby validating the effectiveness of the proposed design. The preservation of physiological periodicity demonstrates the capability of the system to extract vital sign information, even in a multi-subject sensing scenario.

Fig. 12 illustrates the frequency-domain spectra (i.e., (16)) of the extracted physiological signals for each of the three patients. Specifically, Fig. 12(a) presents the spectral representation of the respiration signals. The ground truth respiration frequencies are $f_{r,1} = 0.25$ Hz, $f_{r,2} = 0.3145$ Hz, and $f_{r,3} = 0.18$ Hz. As observed, the estimated spectral peaks closely match the true respiration frequencies across all patients, confirming the accuracy of the proposed method. Similarly, Fig. 12(b) shows the spectra of the heartbeat signals,

with corresponding ground truth frequencies of $f_{h,1} = 1.2$ Hz, $f_{h,2} = 1.345$ Hz, and $f_{h,3} = 1.0578$ Hz. The observed spectral peaks align well with these values, further validating the reliability of the proposed design in capturing vital sign information.

Fig. 13 evaluates the impact of beamforming on the accuracy of heart rate and respiration rate estimations. The results clearly demonstrate that beamforming significantly reduces estimation error for both physiological parameters. Specifically, for heart rate estimation, beamforming leads to a nearly 60% reduction in RMSE. A comparable improvement is observed in respiration rate estimation, where beamforming reduces the RMSE by approximately 55%. These findings underscore the effectiveness of beamforming in enhancing the precision of vital sign monitoring.

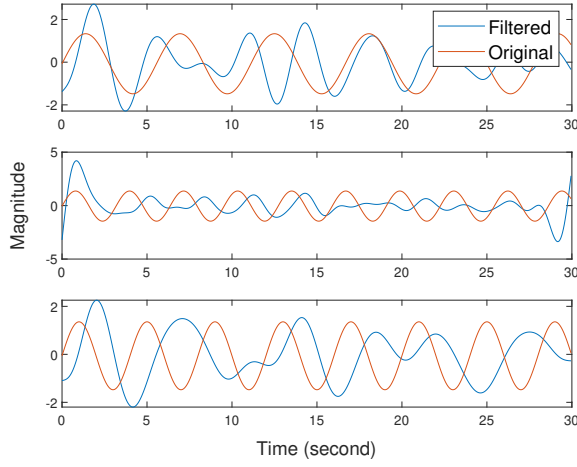
Table III compares the HR and RR estimation performance for single and multi-patient scenarios, evaluated using RMSE. In the single-patient case, Patient 1's ground truth values are 72 BPM and 15 RPM. The benchmark sensing-only design achieves RMSEs of 0.3866 BPM and 0.5028 RPM. The proposed joint design achieves an RMSE of 0.5386 BPM and 0.7749 RPM. While the RR error increases by approximately 54% relative to the sensing-only design, the relative errors remain low (0.75% for HR and 5.17% for RR), indicating that the joint design preserves strong estimation accuracy. In the multi-patient scenario, inter-user interference leads to performance degradation. Using VMD combined with IMFs to separate composite signals, Patient 1 achieves HR and RR RMSEs of 0.8101 BPM and 0.9682 RPM, corresponding to relative errors of 1.13% and 6.45%, respectively. Patient 2, with ground truth values of 80.7 BPM and 18.87 RPM, obtains RMSEs of 1.4867 BPM and 0.6426 RPM (relative errors of 1.84% and 3.41%). Patient 3's HR and RR are 63.47 BPM and 10.80 RPM, with RMSEs of 1.1305 BPM and 1.3107 RPM (relative errors of 1.97% and 12.14%). Compared to the FFT-only method, the VMD & IMFs approach delivers substantial accuracy gains across all patients, highlighting its effectiveness in mitigating multi-user interference. Using the VMD & IMFs method, when the noise variance increases from 10^{-3} to 10^0 , a significant degradation in HR estimation is observed, while RR estimation experiences a smaller performance drop. This difference arises because the amplitude of HR signals is much smaller than that of RR signals, making HR more susceptible to environmental noise. In contrast, the FFT-based method is less sensitive to noise variations but yields substantially lower estimation accuracy overall.

VII. CONCLUSION AND FUTURE DIRECTIONS

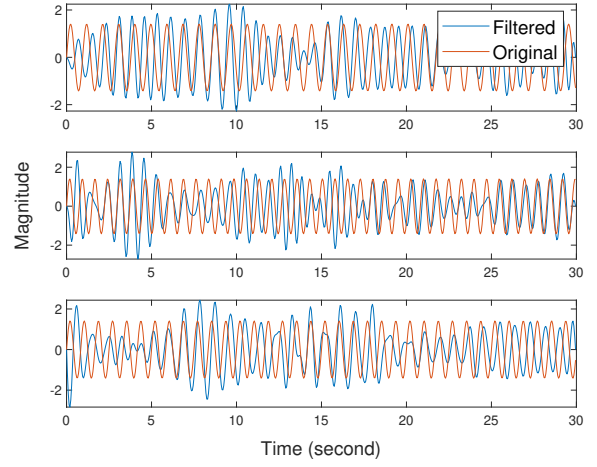
In this paper, we have proposed an integrated framework that combines sensing, computing, and semantic communication for e-health applications, with a particular focus on vital sign monitoring. The system comprises a service robot, multiple healthcare devices, and several patients. The service robot employs IMM filters to track patient movements and design predictive sensing beamformers accordingly. To enhance data efficiency and protect privacy, we have applied semantic extraction to select representative information from the sensing

Table II: Average tracking performance across time for varying noise, speed, and patient movement probabilities.

Noise σ_e^2	Speed (m/s)	Probability of each movement	RMSE $_{\theta,L}$ (degree)	RMSE $_{d,L}$ (m)
10^0	0.1	0.05	0.26843	0.18046
		0.15	0.26856	0.18434
		0.25	0.26968	0.18737
	1	0.05	0.26843	0.18689
		0.15	0.26856	0.19295
		0.25	0.26969	0.19724
	3	0.05	0.26843	0.22393
		0.15	0.26859	0.23551
		0.25	0.26973	0.24438
10^{-3}	0.1	0.05	0.011145	0.011112
		0.15	0.011151	0.011118
		0.25	0.011175	0.011122
	1	0.05	0.011145	0.011118
		0.15	0.011151	0.011126
		0.25	0.011175	0.011132
	3	0.05	0.011145	0.011141
		0.15	0.011151	0.011142
		0.25	0.011175	0.011149

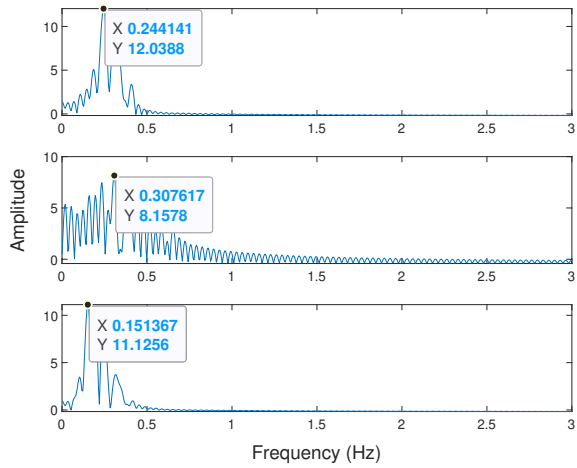


(a) Respiration signal phase over time for patients 1–3.

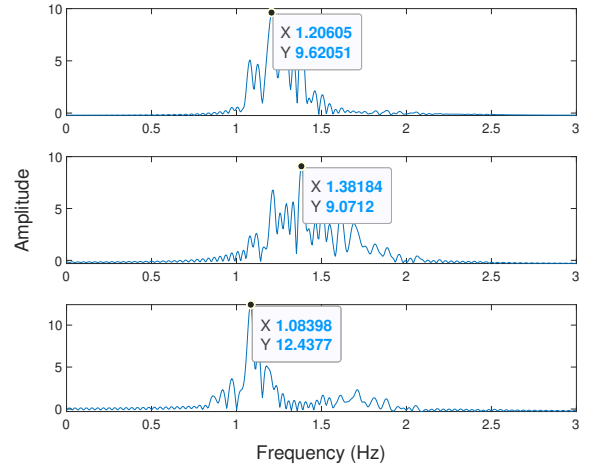


(b) Heartbeat signal phase over time for patients 1–3.

Figure 11: Temporal signal phase variations.



(a) Frequency spectrum of respiration signals for patients 1–3.



(b) Frequency spectrum of heartbeat signals for patients 1–3.

Figure 12: Frequency-domain analysis.

Table III: Heart rate and respiration rate estimation error with single patient and multiple patients.

	Noise σ_e^2	Patient Number	RMSE _{h,l} (BPM)	RMSE _{r,l} (RPM)
Single Patient (Joint design)	10^{-3}	Patient 1	0.538553	0.774975
Single Patient (Sensing-only design)	10^{-3}	Patient 1	0.386541	0.502827
Multiple Patients (Using VMD & IMFs)	10^0	Patient 1	1.344375	0.968246
		Patient 2	1.486659	0.642626
		Patient 3	1.253420	1.445899
Multiple Patients (Using VMD & IMFs)	10^{-3}	Patient 1	0.810093	0.968246
		Patient 2	1.486659	0.642626
		Patient 3	1.130528	1.310713
Multiple Patients (FFT only)	10^0	Patient 1	3.303053	2.432688
		Patient 2	1.486659	3.128573
		Patient 3	4.409326	2.035045
Multiple Patients (FFT only)	10^{-3}	Patient 1	3.303053	2.432688
		Patient 2	1.486659	3.128573
		Patient 3	3.031213	1.310713

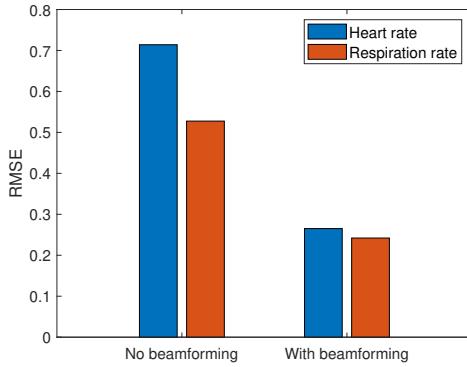


Figure 13: Effect of beamforming on heartbeat and respiration rate estimation.

data. We have evaluated the system performance using key metrics such as semantic computing power and the CRB. We have formulated a joint optimisation problem involving beamforming and semantic extraction ratio, addressing its non-convex nature through a combination of bounding techniques and the bisection method. Simulation results have demonstrated that the proposed framework and algorithms outperform conventional joint sensing and communication methods by achieving higher sensing accuracy, improved semantic transmission efficiency, and enhanced privacy preservation.

As future directions, hardware implementation of the proposed system represents a promising path for practical eHealth applications. Incorporating machine learning techniques can further enhance the robustness of vital sign extraction in noisy and reflective environments. Additionally, integrating complementary sensing modalities, such as cameras, may provide more benefits to system performance. Finally, evaluating the proposed approach within multi-robot or cooperative robotic systems would broaden its applicability and impact in eHealth scenarios.

APPENDIX A PROOF OF PROPOSITION 1

Through considering the Lagrange multiplier η , problem (65) can be transferred to

$$\max_{\rho_k} \sum_{k=1}^K \frac{\iota}{\rho_k} D_k + \eta (F \sum_{k=1}^K \ln \rho_k - P_{c\&s} + P_t) \quad (\text{A.1a})$$

$$\text{s.t. } \rho_{LB,k} \leq \rho_k \leq \rho_{UB,k}, \forall k. \quad (\text{A.1b})$$

The maximum value of the function can be found by taking the first-order derivative. Since this is a multi-variable function, the partial derivative of each $\rho_k, k \in K$ is given by:

$$\frac{\partial}{\partial \rho_k} = \frac{-\iota D_k}{\rho_k^2} + \frac{\eta F}{\rho_k}. \quad (\text{A.2})$$

Therefore, $\rho_k^* = \min \left(\max \left(\frac{\iota D_k}{\eta F}, \rho_{LB,k} \right), \rho_{UB,k} \right)$ is found when $\frac{\partial}{\partial \rho_k} = 0$. When the optimal value of η , denoted by η^* , is found, the optimal value of ρ is also found. The optimal value of η , denoted by η^* , can be found by using the bisection method.

REFERENCES

- [1] X. Li, J. W. You, Z. Gu, Q. Ma, L. Chen, J. Zhang, S. Jin, and T. J. Cui, "Passive human sensing enhanced by reconfigurable intelligent surface: Opportunities and challenges," *IEEE Communications Magazine*, 2024.
- [2] Q. Zhai, X. Han, Y. Han, J. Yi, S. Wang, and T. Liu, "A contactless on-bed radar system for human respiration monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–10, 2022.
- [3] D. Zhang, X. Zhang, Y. Xie, F. Zhang, H. Yang, and D. Zhang, "From single-point to multi-point reflection modeling: Robust vital signs monitoring via mmwave sensing," *IEEE Transactions on Mobile Computing*, 2024.
- [4] F. Wang, X. Zeng, C. Wu, B. Wang, and K. R. Liu, "mmhrv: Contactless heart rate variability monitoring using millimeter-wave radio," *IEEE Internet of Things Journal*, vol. 8, no. 22, pp. 16623–16636, 2021.
- [5] J. Liu, J. Wang, Q. Gao, X. Li, M. Pan, and Y. Fang, "Diversity-enhanced robust device-free vital signs monitoring using mmwave signals," *IEEE Transactions on Mobile Computing*, 2024.
- [6] F. Fioranelli and J. L. Kurnec, "Contactless radar sensing for health monitoring," *Engineering and Technology for Healthcare*, pp. 29–59, 2021.
- [7] D. Manimegalai and M. Karthikeyan, "Exploring the impact of 5g and the promise of 6g: Transforming healthcare strategies and technology," in *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*, pp. 1–8, IEEE, 2023.

- [8] A. Dolas, R. Bodile, A. Kaushik, A. Kaur, R. Singh, and P. Chatzimisios, "Integrated ai and 6g driven e-health: Enabling design, challenges, and future prospects," in *2024 IEEE Conference on Standards for Communications and Networking (CSCN)*, pp. 371–376, IEEE, 2024.
- [9] D. Wen, Y. Zhou, X. Li, Y. Shi, K. Huang, and K. B. Letaief, "A survey on integrated sensing, communication, and computation," *IEEE Communications Surveys & Tutorials*, 2024.
- [10] K. Meng, C. Masouros, A. P. Petropulu, and L. Hanzo, "Cooperative isac networks: Opportunities and challenges," *IEEE Wireless Communications*, 2024.
- [11] X. Zhu, J. Liu, L. Lu, T. Zhang, T. Qiu, C. Wang, and Y. Liu, "Enabling intelligent connectivity: A survey of secure isac in 6g networks," *IEEE Communications Surveys & Tutorials*, 2024.
- [12] Q. Qi, X. Chen, and C. Yuen, "Joint offloading selection and resource allocation for integrated localization and computing in edge-intelligent networks," *IEEE Transactions on Vehicular Technology*, 2024.
- [13] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Towards dual-functional wireless networks for 6g and beyond," *IEEE journal on selected areas in communications*, 2022.
- [14] Q. Qi, X. Chen, A. Khalili, C. Zhong, Z. Zhang, and D. W. K. Ng, "Integrating sensing, computing, and communication in 6g wireless networks: Design and optimization," *IEEE Transactions on Communications*, vol. 70, no. 9, pp. 6212–6227, 2022.
- [15] Q. Qi, X. Chen, C. Zhong, and Z. Zhang, "Integrated sensing, computation and communication in b5g cellular internet of things," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 332–344, 2020.
- [16] A. Olfat and M. Shikh-Bahaei, "Optimum power and rate adaptation for mqam in rayleigh flat fading with imperfect channel estimation," *IEEE transactions on vehicular technology*, vol. 57, no. 4, pp. 2622–2627, 2008.
- [17] H. Bobarshad, M. van der Schaar, and M. R. Shikh-Bahaei, "A low-complexity analytical modeling for cross-layer adaptive error protection in video over wlan," *IEEE Transactions on Multimedia*, vol. 12, no. 5, pp. 427–438, 2010.
- [18] K. Nehra, A. Shadmand, and M. Shikh-Bahaei, "Cross-layer design for interference-limited spectrum sharing systems," in *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pp. 1–5, IEEE, 2010.
- [19] K. Nehra and M. Shikh-Bahaei, "Spectral efficiency of adaptive mqam/ofdm systems with cfo over fading channels," *IEEE transactions on vehicular technology*, vol. 60, no. 3, pp. 1240–1247, 2010.
- [20] G. Jia, Z. Yang, H.-K. Lam, J. Shi, and M. Shikh-Bahaei, "Channel assignment in uplink wireless communication using machine learning approach," *IEEE Communications Letters*, vol. 24, no. 4, pp. 787–791, 2020.
- [21] A. Kobrahi and M. Shikh-Bahaei, "Cross-layer adaptive arq and modulation tradeoffs," in *2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–5, IEEE, 2007.
- [22] N. Pathak, A. Mukherjee, and S. Misra, "Sembox: Semantic interoperability in a box for wearable e-health devices," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 5, pp. 2306–2313, 2022.
- [23] S. Long, J. Tan, B. Mao, F. Tang, Y. Li, M. Zhao, and N. Kato, "A survey on intelligent network operations and performance optimization based on large language models," *IEEE Communications Surveys & Tutorials*, 2025.
- [24] Z. Guo, F. Tang, L. Luo, M. Zhao, and N. Kato, "A survey on applications of large language model-driven digital twins for intelligent network optimization," *IEEE Communications Surveys & Tutorials*, 2025.
- [25] Z. Yang, M. Chen, Z. Zhang, and C. Huang, "Energy efficient semantic communication over wireless networks with rate splitting," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1484–1495, 2023.
- [26] Z. Yang, M. Chen, G. Li, Y. Yang, and Z. Zhang, "Secure semantic communications: Fundamentals and challenges," *arXiv preprint arXiv:2301.01421*, 2023.
- [27] J. Wang, Y. Yang, Z. Yang, C. Huang, M. Chen, Z. Zhang, and M. Shikh-Bahaei, "Generative ai empowered semantic feature multiple access (sfma) over wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, 2025.
- [28] Y. Yang, J. Zhou, Z. Yang, and M. Shikh-Bahaei, "Fluid antenna-enabled near-field integrated sensing, computing and semantic communication for emerging applications," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2025.
- [29] M. A. Rahman and S. Al-Hazzaa, "Next-generation virtual hospital: Integrating discriminative and large multi-modal generative ai for personalized healthcare," in *GLOBECOM 2024-2024 IEEE Global Communications Conference*, pp. 3509–3514, IEEE, 2024.
- [30] X. Peng, Z. Qin, X. Tao, J. Lu, and L. Hanzo, "A robust semantic text communication system," *IEEE Transactions on Wireless Communications*, vol. 23, no. 9, pp. 11372–11385, 2024.
- [31] N. G. Evgenidis, N. A. Mitsiou, S. A. Tegos, P. D. Diamantoulakis, P. Sarigiannidis, I. Krikidis, and G. K. Karagiannidis, "Hybrid semantic-shannon communications," *IEEE Transactions on Wireless Communications*, vol. 23, no. 9, pp. 10926–10940, 2024.
- [32] Z. Hu, C. You, T. Liu, D. Wen, Y. Hu, Y. Cui, Y. Gong, and K. Huang, "Semantic communication meets edge intelligence: Semantic-relay-aided text transmissions," *IEEE Internet of Things Journal*, 2024.
- [33] A. Madhavi, A. H. Shnain, B. Gayathri, et al., "Iot-enabled smart system for personalized health management with semantic technologies," in *2025 International Conference on Intelligent Control, Computing and Communications (IC3)*, pp. 1153–1158, IEEE, 2025.
- [34] S. N. Karahan and O. Kaya, "Towards 6g semantic communication: Technologies, applications, and research challenges," in *2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (ICHORA)*, pp. 1–7, IEEE, 2025.
- [35] C. Yuan, H. Ye, and Y. Miao, "Universal image semantic communication for edge network," in *2024 4th International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*, pp. 316–320, IEEE, 2024.
- [36] Y. Wang, Z. Wang, J. A. Zhang, H. Zhang, and M. Xu, "Vital sign monitoring in dynamic environment via mmwave radar and camera fusion," *IEEE Transactions on Mobile Computing*, vol. 23, no. 5, pp. 4163–4180, 2023.
- [37] A. Athama, K. Wang, X. Chen, and Y. Li, "Semantic communications for healthcare applications: opportunities and challenges," in *2024 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT)*, pp. 383–388, IEEE, 2024.
- [38] Q. Fan, Y. Xie, C. Zhang, X. Liu, and L. Zhu, "An authentic and privacy-preserving scheme towards e-health data transmission service," *IEEE Transactions on Services Computing*, 2024.
- [39] S. T. Ahmed, A. Kaladevi, A. Shankar, F. Alqahtani, et al., "Privacy enhanced edge-ai healthcare devices authentication: a federated learning approach," *IEEE Transactions on Consumer Electronics*, 2025.
- [40] A. B. Mpembele, T. Rogers, U. Ghosh, and S. Shetty, "Communication-efficient and privacy-preserving edge-cloud framework for smart healthcare," in *2023 IEEE Globecom Workshops (GC Wkshps)*, pp. 377–382, IEEE, 2023.
- [41] J. Romeo, M. Abbass, A. Sherif, M. M. K. Mamun, M. Elersy, and K. Khalil, "Privacy-preserving machine learning for e-health applications: A survey," in *2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI)*, pp. 1–6, IEEE, 2024.
- [42] N. Azam, A. L. Michala, S. Ansari, and N. B. Truong, "Modelling technique for gdpr-compliance: Toward a comprehensive solution," in *GLOBECOM 2023-2023 IEEE global communications conference*, pp. 3300–3305, IEEE, 2023.
- [43] E. Mazar, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: a survey," *IEEE Transactions on aerospace and electronic systems*, vol. 34, no. 1, pp. 103–123, 2002.
- [44] C. Jiang, C. Zhang, C. Huang, J. Ge, D. Niyato, and C. Yuen, "Risk-assisted isac systems for robust secure transmission with imperfect sense estimation," *IEEE Transactions on Wireless Communications*, 2025.
- [45] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, and Y. C. Eldar, "Joint transmit beamforming for multiuser mimo communications and mimo radar," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3929–3944, 2020.
- [46] M. Shikh-Bahaei, "Joint optimization of "transmission rate" and "outer-loop snr target" adaptation over fading channels," *IEEE Transactions on Communications*, vol. 55, no. 3, pp. 398–403, 2007.
- [47] F. Wang, F. Zhang, C. Wu, B. Wang, and K. R. Liu, "Vimo: Multiperson vital sign monitoring using commodity millimeter-wave radio," *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 1294–1307, 2020.
- [48] L. Liu, J. Zhang, Y. Qu, S. Zhang, and W. Xiao, "mmrh: Noncontact vital sign detection with an fmcw mm-wave radar," *IEEE Sensors Journal*, vol. 23, no. 8, pp. 8856–8866, 2023.
- [49] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3834–3862, 2020.
- [50] X. Li, J. Zhang, Z. Cai, X. W. Wu, Q. Peng, Q. Ma, J. W. You, and T. J. Cui, "Dynamic electromagnetic model to detect human vital

- signs based on time-domain finite integration theorem,” *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, 2024.
- [51] M. N. I. Shuzan, M. H. Chowdhury, M. S. Hossain, M. E. Chowdhury, M. B. I. Reaz, M. M. Uddin, A. Khandakar, Z. B. Mahbub, and S. H. M. Ali, “A novel non-invasive estimation of respiration rate from motion corrupted photoplethysmograph signal using machine learning model,” *IEEE Access*, vol. 9, pp. 96775–96790, 2021.
 - [52] L. Zhang, X. Lei, T. Ma, H. Niu, and C. Yuen, “Joint user localization, channel estimation, and pilot optimization for ris-isac,” *IEEE Transactions on Wireless Communications*, 2024.
 - [53] Y. Xiang, Y. Gao, X. Yang, S. Kang, and M. Shao, “An esprint-based moving target sensing method for mimo-ofdm isac systems,” *IEEE Communications Letters*, vol. 27, no. 12, pp. 3205–3209, 2023.
 - [54] F. Liu, W. Yuan, C. Masouros, and J. Yuan, “Radar-assisted predictive beamforming for vehicular links: Communication served by sensing,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7704–7719, 2020.
 - [55] X. Meng, F. Liu, C. Masouros, W. Yuan, Q. Zhang, and Z. Feng, “Vehicular connectivity on complex trajectories: Roadway-geometry aware isac beam-tracking,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 7408–7423, 2023.
 - [56] Y. Yuan, M. Wen, X. Zheng, S. Wang, and S. Gao, “Predictive target-to-user association in complex scenarios via hybrid-field isac signaling,” *arXiv preprint arXiv:2501.10676*, 2025.
 - [57] T. Kirubarajan and Y. Bar-Shalom, “Kalman filter versus imm estimator: when do we need the latter?,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1452–1457, 2003.
 - [58] Z. Xu, S. Xu, H. Ding, and R. Xu, “An isac-based beam tracking scheme against inter-region interference for the multi-rsu v2i scenario,” *IEEE Transactions on Vehicular Technology*, 2024.
 - [59] K. Jo, K. Chu, K. Lee, and M. Sunwoo, “Integration of multiple vehicle models with an imm filter for vehicle localization,” in *2010 IEEE intelligent vehicles symposium*, pp. 746–751, IEEE, 2010.
 - [60] N. Nadarajah, R. Tharmarasa, M. McDonald, and T. Kirubarajan, “Imm forward filtering and backward smoothing for maneuvering target tracking,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2673–2678, 2012.
 - [61] W. J. Farrell, “Interacting multiple model filter for tactical ballistic missile tracking,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 2, pp. 418–426, 2008.
 - [62] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.
 - [63] Z. Zhao, Z. Yang, Y. Hu, C. Zhu, M. Shikh-Bahaei, W. Xu, Z. Zhang, and K. Huang, “Compression ratio allocation for probabilistic semantic communication with rsma,” *IEEE Transactions on Communications*, 2025.
 - [64] Y. Yang, M. Shikh-Bahaei, Z. Yang, C. Huang, W. Xu, and Z. Zhang, “Secure design for integrated sensing and semantic communication system,” in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–7, IEEE, 2024.
 - [65] Y. Yang, Z. Yang, C. Huang, W. Xu, Z. Zhang, D. Niyato, and M. Shikh-Bahaei, “Integrated sensing, computing and semantic communication for vehicular networks,” *IEEE Transactions on Vehicular Technology*, pp. 1–6, 2025.
 - [66] W. Wei, X. Pang, X. Qin, S. Gong, C. Xing, N. Zhao, and D. Niyato, “Cramér-rao bound and secure transmission trade-off design for semi-rs-enabled isac,” *IEEE Transactions on Wireless Communications*, 2024.
 - [67] I. Bekkerman and J. Tabrikian, “Target detection and localization using mimo radars and sonars,” *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3873–3883, 2006.
 - [68] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, “Outage constrained robust transmit optimization for multiuser mimo downlinks: Tractable approximations by conic optimization,” *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5690–5705, 2014.
 - [69] F. Wang, X. Wang, and Y. Zhu, “Transmit beamforming for multiuser downlink with per-antenna power constraints,” in *2014 IEEE International Conference on Communications (ICC)*, pp. 4692–4697, IEEE, 2014.
 - [70] E. Song, Q. Shi, M. Sanjabi, R.-Y. Sun, and Z.-Q. Luo, “Robust sinr-constrained mimo downlink beamforming: When is semidefinite programming relaxation tight?,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, pp. 1–11, 2012.
 - [71] M. B. Shenouda and T. N. Davidson, “Convex conic formulations of robust downlink precoder designs with quality of service constraints,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 714–724, 2007.
 - [72] P. Zhao, M. Zhang, H. Yu, H. Luo, and W. Chen, “Robust beamforming design for sum secrecy rate optimization in mu-miso networks,” *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1812–1823, 2015.
 - [73] M. Grant and S. Boyd, “Cvx: Matlab software for disciplined convex programming, version 2.1,” 2014.
 - [74] Z. Chen, F. Wang, G. Han, X. Wang, and V. K. Lau, “Robust beamforming design for secure near-field isac systems,” *IEEE Wireless Communications Letters*, 2025.
 - [75] Y. Chen, Y. Feng, X. Li, L. Zhao, and K. Shen, “Fast fractional programming for multi-cell integrated sensing and communications,” *IEEE Transactions on Wireless Communications*, 2025.