




The Short-spacing Interferometer Array for Global 21 cm Signal Detection (SIGMA): Design of the Antennas and Layout

Feiyu Zhao^{1,2}, Quan Guo^{1,3}, Qian Zheng^{1,3}, Ruxi Liang¹, Pengfei Zhang⁴, Yajun Wu¹, Junhua Gu⁵ , Zhao Yang⁴, Yun Yu¹, Yan Huang⁵, and Tianyang Liu^{1,2}

¹ Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China; guoquan@shao.ac.cn, qzheng@shao.ac.cn

² University of Chinese Academy of Sciences, Beijing 101408, China

³ Key Laboratory of Radio Astronomy and Technology, Chinese Academy of Sciences, Beijing 100101, China

⁴ Xidian University and National Key Laboratory of Radar Detection and Sensing, Xi'an 710071, China

⁵ National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; jhgu@nao.cas.cn

Received 2024 November 25; revised 2025 March 25; accepted 2025 March 27; published 2025 May 9

Abstract

Numerous experiments have been designed to investigate the Cosmic Dawn (CD) and Epoch of Reionization (EoR) by examining redshifted 21 cm emissions from neutral hydrogen. Detecting the global spectrum of redshifted 21 cm signals is typically achieved through single-antenna experiments. However, this global 21 cm signal is deeply embedded in foreground emissions, which are about four orders of magnitude stronger. Extracting this faint signal is a significant challenge, requiring highly precise instrumental calibration. Additionally, accurately modelling receiver noise in single-antenna experiments is inherently complex. An alternative approach using a short-spacing interferometer is expected to alleviate these difficulties because the noise in different receivers is uncorrelated and averages to zero upon cross-correlation. The Short-spacing Interferometer Array for Global 21 cm Signal Detection (SIGMA) is an upcoming experiment aimed at detecting the global CD/EoR signal using this approach. We describe the SIGMA system with a focus on optimal antenna design and layout, and propose a framework to address cross-talk between antennas in future calibrations. The SIGMA system is intended to serve as a prototype to gain a better understanding of the system's instrumental effects and to optimize its performance further.

Key words: General – instrumentation: interferometers – radio continuum: galaxies – methods: observational

1. Introduction

The redshifted 21 cm signal of neutral hydrogen from the Cosmic Dawn (CD) and Epoch of Reionization (EoR) is detectable in the low-frequency radio band, ranging from 50 to 200 MHz. Theoretical models suggest that this signal is extremely weak, overshadowed by foreground emissions that are 4–5 orders of magnitude higher (Furlanetto et al. 2006). Several radio telescopes have been constructed with the detection of the CD/EoR signal as a primary scientific goal. Methods for detecting CD/EoR include: (a) statistical analysis of CD/EoR structures using radio interferometers; (b) tomographic approaches with radio interferometers; and (c) observation of the global CD/EoR signal using single antennas. Challenges in these detections involve removing foreground interference, achieving high-precision instrument calibration, and ensuring a high signal-to-noise ratio in both observation and data analysis for interferometers and single-antenna experiments.

The Square Kilometer Array (SKA, Koopmans et al. 2015) is set to become the world's largest ground-based radio telescope, with CD/EoR detection as one of its key projects.

SKA plans to conduct deep imaging of five specific sky fields and to statistically measure CD/EoR structures through extensive field surveys. SKA pathfinders, such as the Murchison Widefield Array (MWA, Barry et al. 2019b), Low Frequency Array (LOFAR, Mertens et al. 2020), Giant Metrewave Radio Telescope (Paciga et al. 2013), and 21CMA (Zheng et al. 2016), focus on statistical measurements of CD/EoR. However, due to sensitivity constraints, these pathfinders are currently unable to perform tomographic measurements of CD/EoR. Significant work in statistical CD/EoR measurements is still required (Dilullo et al. 2020; Hothi et al. 2021; Kolopanis et al. 2023; Munshi et al. 2023; Shaw et al. 2023), with numerous challenges remaining in data reduction and observational techniques.

Meanwhile, numerous experiments have been conducted to detect the global CD/EoR signal using individual antennas. Some of these experiments include EDGES (Bowman et al. 2008), SCI-HI (Voytek et al. 2014), BIGHORNS (Sokolowski et al. 2015), LEDA (Bernardi et al. 2015), SARAS (Patra et al. 2013), PRIZM (Philip et al. 2019), and REACH (de Lera Acedo et al. 2022). In a study by Bowman et al. (2018), a potential CD signal feature was reported by EDGES. This

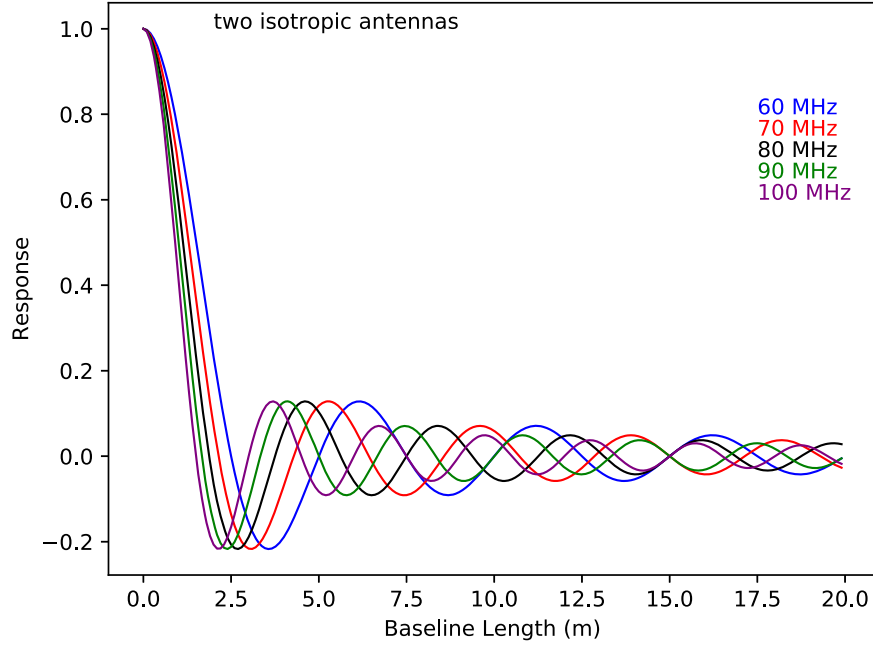


Figure 1. Responses of two isotropic short spacing antennas at different frequencies.

feature is characterized by an absorption feature at 78.2 MHz with an amplitude of approximately 556 mK. However, these findings, which challenge existing physical models, have recently been contested by results from the SARAS3 experiment (Bevins et al. 2022). Further corroboration of these findings is needed through additional experimental observations.

In addition to the detection of the global signal from a single antenna, Vedantham et al. (2015) demonstrated a method for detecting the global redshifted 21 cm signal using interferometry, noting that short-spaced interferometers respond significantly to the global signal. Presley et al. (2015) discussed designing a global signal interferometer and performed simulations to evaluate its performance. Zhang et al. (2023) explored the possibility of extracting the 21 cm CD signal from measurements of the interferometer array, focusing on the feasibility of recovering the global spectrum from visibility data. Singh et al. (2015) developed the theoretical framework for interferometer responses to global signals using dipole and resonant loop antennas, estimating their sensitivity. McKinley et al. (2020) utilized an array of short-spacing antennas to detect the global CD signal interferometrically, asserting the theoretical feasibility of extracting the global redshifted 21 cm signal using this method. Thekkeppattu et al. (2022) reported on the Short Spacing Interferometer Telescope (SITARA), a prototype designed to measure global CD/EoR signals. Their further studies (Thekkeppattu et al. 2023) delve into calibration, data analysis methods, and various factors influencing the performance of SITARA.

In this study, we present an experimental project, which is planned to be established in China, named the Short-spacing Interferometer Array for Global 21 cm Signal Detection (SIGMA). This initiative is designed to detect the global H I 21 cm signal using an interferometric approach with short-spacing antennas. A key aspect of our design involves addressing the challenge of mutual antenna coupling, a significant consideration in short-spacing antenna configurations as highlighted by McKinley et al. (2020).

The structure of this paper is outlined as follows: Section 2 introduces the concept of interferometrically measuring the global CD/EoR signal. In Section 3, we provide detailed information on the design of the antennas and their layout. Site selection is also briefly introduced in this section, along with an overview of the electronic design of SIGMA. Section 4 delves into the effect of crosstalk. Section 5 describes data processing. Finally, Section 6 presents discussions and conclusions.

2. Detecting the Global Signal with Closely Spacing Interferometer

Global 21 cm signal can be measured with an interferometer when baselines are short enough to have a response to the global signal. The response of a two-element interferometer can be written as Thompson et al. (2017):

$$V(\mathbf{b}, \nu) = \frac{1}{4\pi} T_{\text{sky}}(\nu) \int A(\mathbf{r}, \nu) e^{-2\pi i(\mathbf{b} \cdot \mathbf{r} / \lambda)} d\Omega, \quad (1)$$

where V is the visibility in the units of K . \mathbf{b} is the baseline vector and ν is the observation frequency. T_{sky} is the brightness



Figure 2. Location of candidate sites near the city of Shanshan.

temperature of the sky. r is the unit vector of the sky direction. A is the beam pattern of the antenna. λ is the wavelength. Ω is the solid angle. Figure 1, similar to Figure 1 in Vedantham et al. (2015), shows that the interferometer's response to a global signal exhibits a *sinc* shape as a function of baseline length across several different frequencies. When the spacing of the antenna elements is less than half the wavelength, there is a significant interferometric response to the global signal.

As described in McKinley et al. (2020), for real observation, the sky brightness temperature is not uniform but consists of a global component and angular variations. Consider $T_{\text{angular}}(\mathbf{r}, \nu) = T_{\text{sky}}(\mathbf{r}, \nu) - T_{\text{sky}}(\nu)$ and the visibility equation can be written as:

$$\begin{aligned} V(\mathbf{b}, \nu) &= \frac{1}{4\pi} (T_{\text{sky}}(\nu) \int A(\mathbf{r}, \nu) e^{-2\pi i(\mathbf{b}\cdot\mathbf{r}/\lambda)} d\Omega \\ &\quad + \int T_{\text{angular}}(\mathbf{r}, \nu) A(\mathbf{r}, \nu) e^{-2\pi i(\mathbf{b}\cdot\mathbf{r}/\lambda)} d\Omega) \\ &= T_{\text{sky}}(\nu) \times \text{global response} + \text{angular response}, \end{aligned} \quad (2)$$

Short baselines respond better to large-scale structures and are insensitive to small angles, and we can assume that the first term of the equation dominates visibility. The $T_{\text{sky}}(\nu)$ can then be estimated from the global response and visibility.

3. Experiment Design

3.1. Site Selection

The experiment requires an RFI-quiet location that balances several considerations: ease of regular access for construction and resupply, and ease of power supply for the data processing center from a nearby power grid. Several potential sites have been surveyed in northwest China, including Qinghai, Gansu, and Xinjiang provinces. One of the best candidates is near the city of Shanshan (E90°12'51."71, N42°51'38."84) in Xinjiang

province (Figure 2). This site is about three hours from the nearest city. We measure the level of radio frequency interference (RFI) around the site twice by surveying with a dipole antenna optimized for the frequency range 50 ~ 200 MHz, two stages of amplifiers used in the SIGMA signal chain, and a Keysight FieldFox RF analyzer N9913A, which means that the level of RFI was amplified by approximately 60 dB before being recorded by the spectrum analyzer. Figure 3 shows the radio spectrum between 50 and 200 MHz, amplified as mentioned, with bandpass filters covering the frequency range of 50–200 MHz. The spectrum has a bin size of 15 kHz, and the data is averaged over ten measurements using the average trace detector. The results suggest that the candidate site is an excellent and quiet location for low-frequency radio observation.

3.2. Antenna Array Design

The redshifted 21 cm signal from the CD/EoR is observable within the 50–200 MHz frequency range, corresponding to wavelengths between 6 m and 1.5 m. For our interferometer to effectively capture the global signal, the baselines must be shorter than half the wavelength, equating to less than 3 m at 50 MHz. According to the findings of Bowman et al. (2018), the absorption valley is located near 78 MHz. Therefore, SIGMA will focus on the 65–90 MHz band in the current phase.

Singh et al. (2015) proposed that a parallel configuration would yield enhanced sensitivity in antenna design. Here we also consider a parallel configuration 1D array distributed along the E–W direction, as depicted in Figure 4. This configuration consists of 20 antennas aligned in a straight line, oriented E–W, with a uniform distance of 1 m between each antenna.

Here, we choose a one-dimensional (1D) array for the following reasons. First, placing all antennas along a single

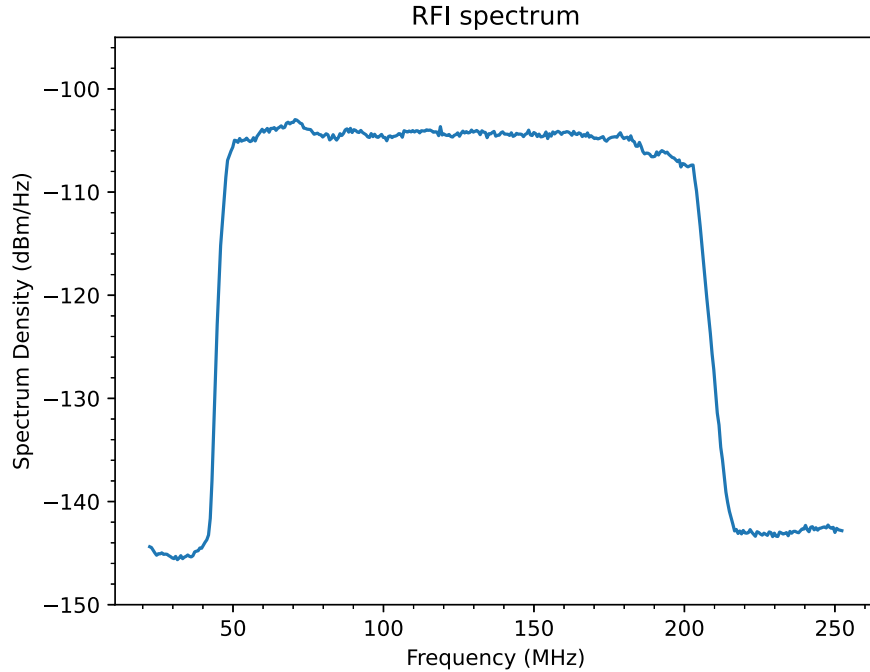


Figure 3. The measured RFI spectrum was amplified by approximately 60 dB with the first and second stages of amplifiers along bandpass filters covering the frequency range of 50–200 MHz used in the SIGMA system.

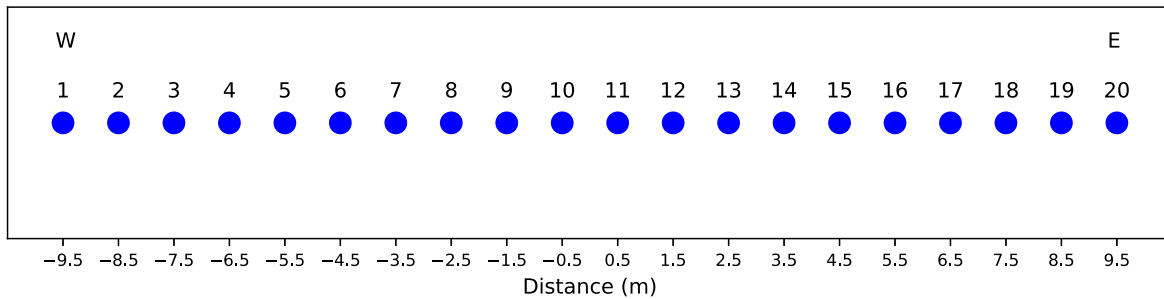


Figure 4. The figure shows the layout of the SIGMA array. The array is arranged east–west, each blue dot corresponds to an antenna, the antennas are 1 m apart, and the numbers on the blue dots are the numbers of the antennas.

baseline helps ensure that each element is in a similar environment, which simplifies both the analytical modeling and the calibration of mutual coupling (Section 4). Second, a linear array can be easily expanded or reduced, making it convenient for iterative experiments and allowing us to adjust the number of elements as needed. Third, since our goal is to measure the large-scale (global) signal rather than produce high-resolution images, we do not require the dense uv coverage that often necessitates two-dimensional layouts. Consequently, a 1D arrangement with short baselines meets our needs without introducing the more complex mutual coupling that can arise in multi-directional arrays.

In this work, we present a reference design with 20 antennas spaced 1 m apart along an East–West line (Figure 4). This

number is chosen mainly to illustrate our design principles and to evaluate antenna performance. However, it is not fixed; future experiments may adjust the array size depending on site conditions, resource availability, or sensitivity requirements.

In our systematic approach to designing antenna elements, we used a methodical process to identify appropriate antenna types, calculate relevant parameters, and assess various antenna designs using simulation. The simulated performance of the antenna will be one of the important criteria for our selection. Given our specific requirements and the power of simulation, we initially selected and evaluated the blade antenna. The blade antenna is favored for its simple construction, wide bandwidth, ease of manufacturing, and near-omnidirectional radiation pattern, which makes it particularly advantageous for global

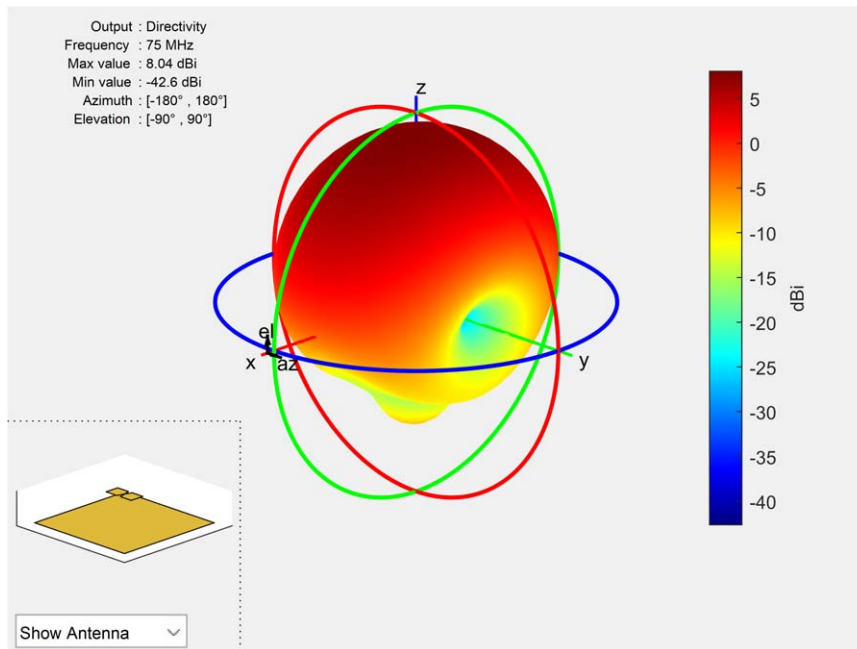


Figure 5. Single blade antenna radiation pattern in 75 MHz.

signal detection. A typical blade antenna element in our design consists of two symmetrical blade panels, each 0.7 m long, 0.65 m wide, and the taper length is 0.6 m. These elements are positioned 1.2 m above the ground, and a 10 m wide, 30 m long metal mesh was placed on the ground. The parameter S_{11} , also known as input return loss, reflects the amount of power that is reflected back from the antenna due to impedance mismatch. A lower value S_{11} is preferable, as it indicates that a greater proportion of the power is effectively transmitted into the antenna, which is crucial for signal capture in our experiment. The radiation pattern and the S_{11} of the isolated single blade antenna are shown in Figures 5 and 6. From the simulation results, it can be seen that the blade antenna performs well in the 65–90 MHz frequency band and is suitable for the SIGMA experiment.

We have modeled the antenna array using the Matlab Antenna Toolbox,⁶ the model is shown in Figure 7. The information about the antenna and array design is summarized in Table 1.

Here we simulated the S_{nn} (input return loss of antenna n) of each antenna of the SIGMA, the termination resistance is internally set to a default of 50 ohms, and the results S_{nn} are shown in Figure 8. S_{nn} is essentially the same except for antenna 1, 20; 2, 19; 3, 18. Compared to the single antenna case, the S_{nn} change significantly due to the effect of mutual coupling, but the antenna is still usable in the 65–90 MHz band.

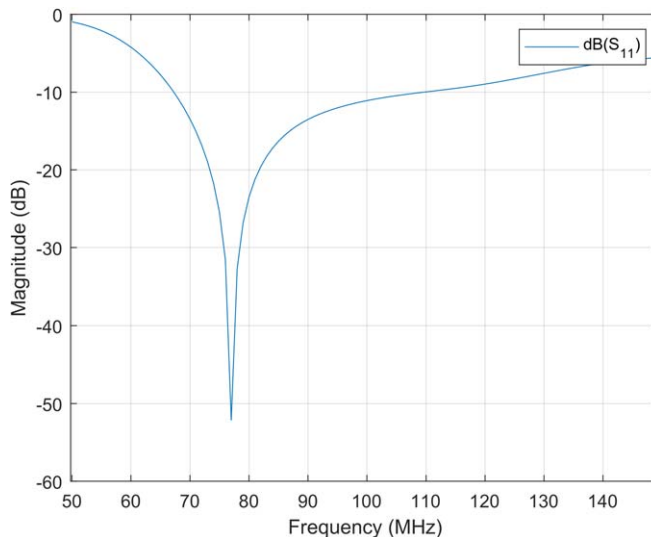


Figure 6. Single blade antenna S_{11} .

Furthermore, the antenna radiation pattern is a critical element that warrants detailed consideration. We simulated the radiation pattern of antenna numbers 1, 4, 7 and 10, the results are shown in Figure 9. The radiation pattern is calculated by driving a specific element in the array. The rest of the array elements are terminated using reference impedance (50 ohms).

In short-spacing interferometer experiments, the antenna spacing is required to be as small as possible in order to obtain as strong a response as possible, which also leads to strong

⁶ <https://ww2.mathworks.cn/products/antenna.html>

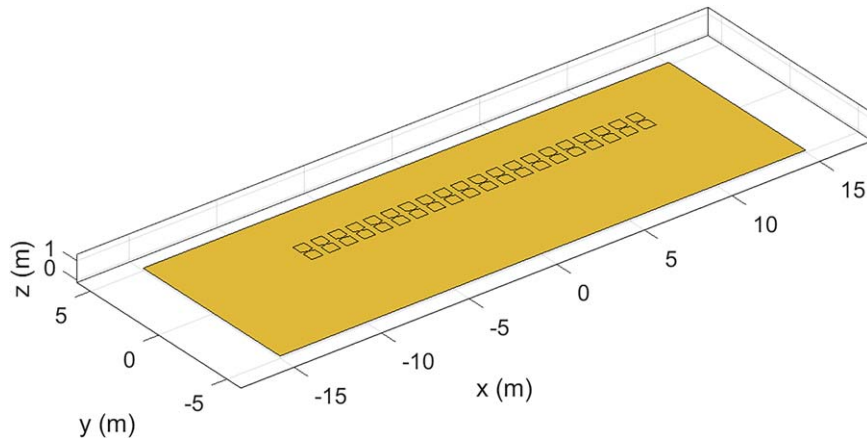


Figure 7. 3D view of the SIGMA 20 blade antennas model. The antennas are in parallel configuration with a height of 1.2 m and the ground plane is 30×10 m of PEC (Perfect Electric Conductor) material.

Table 1
Key Parameters of the SIGMA Design

Parameter	Value	Notes
Frequency range	65–90 MHz	Focus on global EoR signal
Number of antennas	20	1D linear array (E–W direction)
Antenna spacing	1 m	Short baselines for detecting large-scale signal
Antenna type	Blade antenna	Near-omnidirectional, wide bandwidth
Blade size	0.7 m (L), 0.65 m (W), 0.6 m taper	Two symmetric blade panels
Antenna height	1.2 m	Above ground
Ground plane	10 m \times 30 m	Metal mesh (PEC in simulation)

mutual coupling effects. From the simulation results, it can be seen that the S_{nn} and radiation patterns of the antennas change significantly at very close distances, especially for antenna 1, 20; 2, 19; 3, 18, whose performances are more different from those of the other antennas. An approach would be to use these antennas as constraints only, during the actual observation, their observation data will not be used.

The one-dimensional array design allows SIGMA to be easily expanded in the future and higher sensitivity can be obtained by continuing to install more antennas. The scale depends on the condition of the final selected telescope site. The effect of mutual coupling, which will be stronger for shorter baselines, needs to be properly dealt with to measure the global EoR signals. We try to develop a framework in Section 4 and hope to remove it in the calibration. The similar design will be implemented for the high band (90 ~ 200 MHz) in the next phase of SIGMA.

3.3. Electronics Design

In summary, the SIGMA system consists of antennas kept in close proximity, a local “fieldbox” performing analog signal

conditioning, digitization equipped with a time-synchronization client from the White Rabbit (WR) system and a remote correlation center with GPUs performing the receiving and correlation for all antennas. No hardware calibration device such as noise diodes is integrated yet. The initial prototype is designed to be as simple as possible, focusing on studying systematics such as crosstalk between antennas. A block diagram of the SIGMA electronics system is given in Figure 10, and the block diagram of the SIGMA system and data processing is shown in Figure 11.

In the initial prototype system, each antenna is associated with a local fieldbox via a coaxial cable. In the fieldbox, the first stage low noise amplifier (LNA) is assembled based on the chip XT3074 from Chengdu Fairchild Technology Ltd., with a gain of about 27 dB and a noise figure of about 0.5 dB. Through the same coaxial cables the DC power for the first stage LNA is supplied by the fieldbox via bias-tees, which makes the deployment of the fieldbox flexible in terms of location. The second and third stage amplifiers are built based on BFP740ESD from INFINEON, with a gain of about 40 dB and a noise figure of less than 1.2 dB, and ZD3027 from

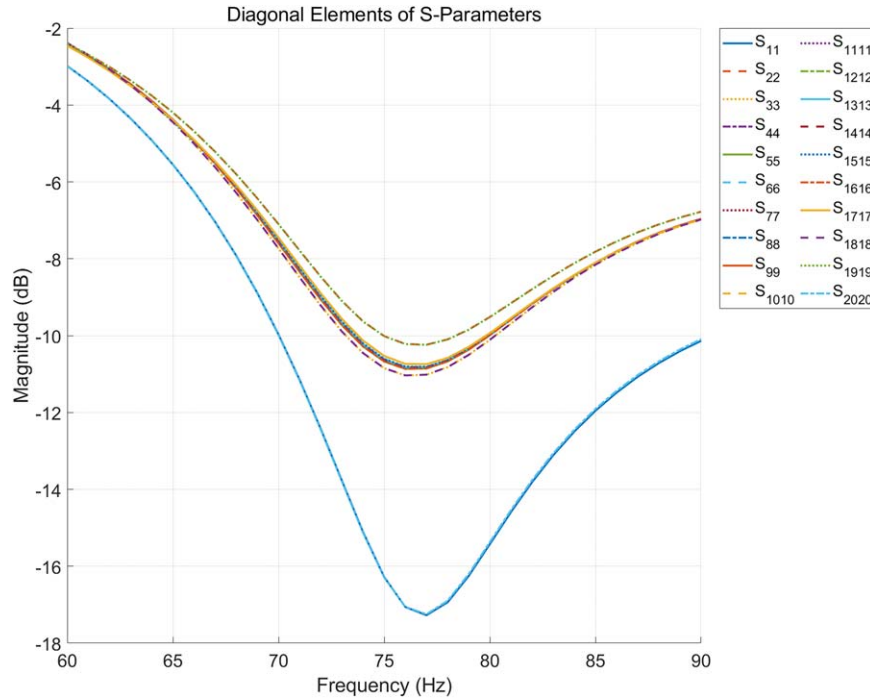


Figure 8. Simulation results of S_{nn} of the SIGMA. Due to the effect of mutual coupling, the S_{nn} are considerably transformed compared to the single antenna case, but they are still usable in 65–90 MHz and especially around 78 MHz. In this figure, S_{11} and S_{2020} overlap (the dark blue line at the bottom and the light blue dashed-dotted line), S_{22} and S_{1919} overlap (the orange-red dashed line at the top and the green dotted line), S_{33} and S_{1818} overlap (the second line from the bottom, shown by the yellow dotted line and the purple dashed line). The remaining traces from S_{44} through S_{1717} essentially overlap one another, indicating that antennas 4 through 17 exhibit nearly identical S_{nn} .

CEmaxRF tech, with a gain of about 25 dB and a noise figure of less than 1.6 dB, respectively. To reduce the effects of out-of-band radio frequency interference (RFI) on signal chain linearity, since the amplifiers used are broadband relative to the required 200 MHz bandwidth, a customized 50–200 MHz bandpass filter with suppression greater than 50 dB outside the bandpass is applied between the amplifier stages and before analog-to-digital conversion (ADC). Together, the analog section has a net gain greater than 80 dB and is shielded by a metal EMC box to reduce interference from outside sources.

The amplified and filtered signals are then digitized in data acquisition nodes, utilizing a customized Synchronized Distributed Array Node Digitizer (SDAND) system. The SDAND system was designed by our team, and further details of the SDAND system will be presented separately (J. Gu et al. 2025, in preparation). This system comprises a set of identical data acquisition units, whose sampling clocks are synchronized using the WR system (Lipiński et al. 2011) with sub-nanosecond accuracy. The SDAND system features several input/output ports, including two coaxial connectors for RF signal input, one optical fiber connector for transporting WR system signals, one optical fiber connector for receiving control messages, and two optical fiber connectors for outputting payload data. The advantages of the SDAND system are

threefold: (1) it enables the formation of a versatile range of radio telescopes, (2) it facilitates the creation of an extensive and flexible data acquisition system with stability, and (3) it transparently handles time synchronization among all data acquisition nodes. Since nearly all types of array-formed radio telescopes rely on synchronously sampling voltage signal outputs from the antenna system, the output data stream of the SDAND system can be appropriately directed to realize various array telescopes, including single beam-forming stations, beam-forming based interferometers (e.g., the SKA LFAA), wide-field monitoring interferometers (e.g., LWA and HERA), pulse-signal triggering and event reconstructing detectors for particle physics studies (e.g., GRAND, Martineau-Huynh et al. 2017), and more. As the signals are transferred in digital mode, data processing centers equipped with high-performance processors and GPUs can be situated far from radio-quiet zones, ensuring sufficient resources for computing and cooling. With the SDAND system’s scalability and flexibility, expanding or repurposing the SIGMA system is relatively straightforward. This can be achieved by adding more antennas with SDAND devices, modifying the software, and increasing the GPU cluster’s data acquisition and correlation computing capabilities as needed.

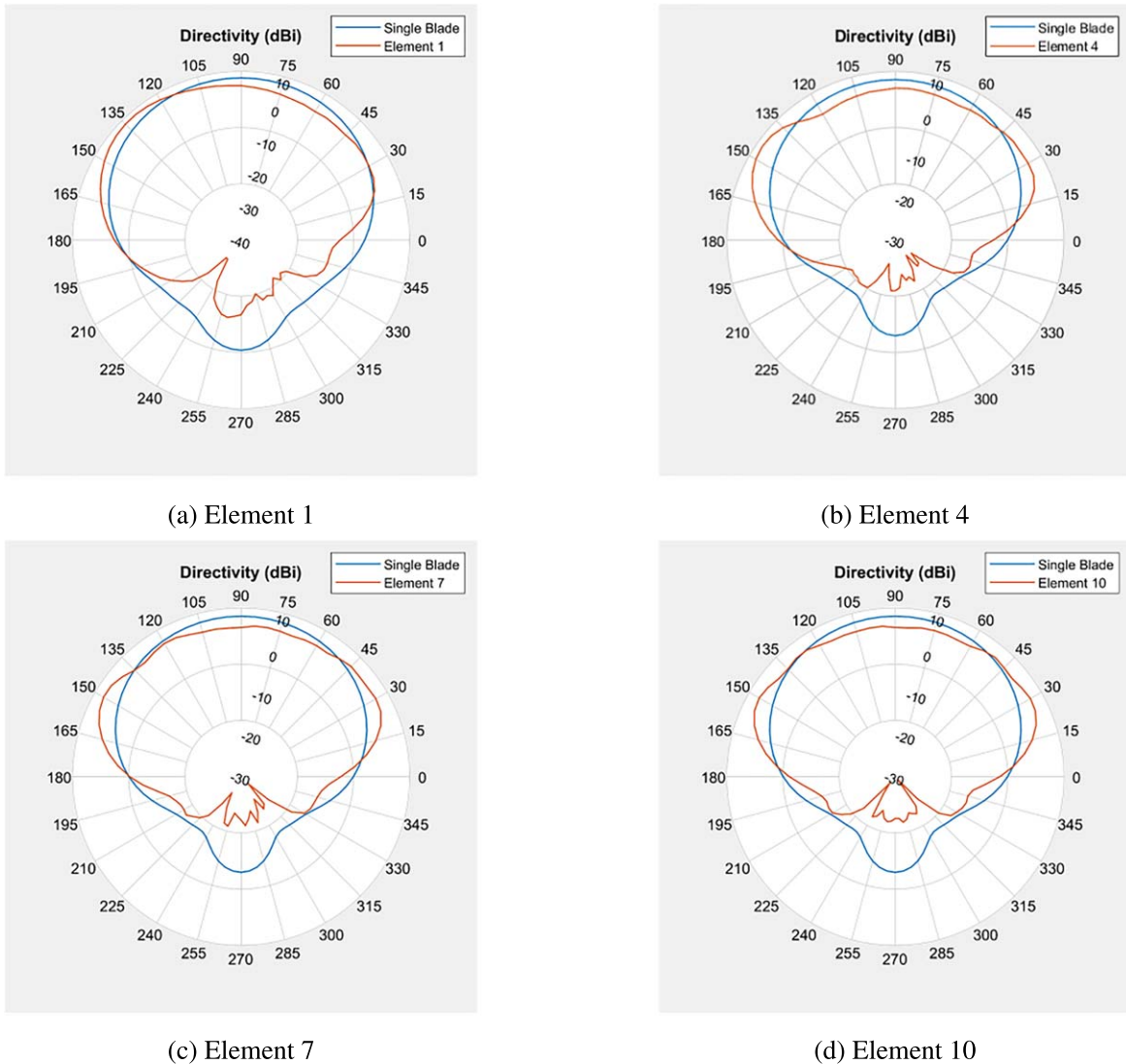


Figure 9. Comparison of the radiation pattern of elements 1, 4, 7, and 10 in the array (EW direction) with single blade antenna in dBi, the top corresponding to the zenith. Antennas in the array still respond well to a wide range. There are some variations in the radiation patterns of the different antennas due to the antenna positions, especially for antenna 1 (symmetrical, antenna 20). In our experiment, radiation patterns below the plane of the antenna (180° – 360°) do not affect observations.

The prototype of the SIGMA system comprises seven data acquisition nodes, each receiving signals from two RF signal processing chains, along with a data processing node equipped with four Nvidia A40 GPU cards and two 100 Gbps Ethernet interface cards. All the data acquisition nodes and the data processing node are connected to a 100 Gbps switch, with the processing node and Ethernet switch located far away from the radio-quiet zone. The data acquisition nodes stream the sampled payload data to the processing node. Real-time channelization and correlation are performed on GPU cards of the data processing node to generate auto- and cross-correlation spectral products. The acquired time domain data

series are channelized into 65,536 spectral channels. Since the data acquisition system is a direct sampling system, each data point is a real number, resulting in only 32,768 independent channels spanning a frequency range of 0–200 MHz, with a spectral resolution of about 6.1 kHz. With the bandpass filter in the RF signal chain, only signals of frequency between 50 and 200 MHz are sensitive to sky signals. All the data acquisition nodes and the data processing node are connected to a 100 Gbps Ethernet switch. The payload data streams from the eight acquisition nodes (i.e., 14 data streams) are routed to the data processing node, where the auto- and cross-correlation products are computed. The computing tasks are evenly

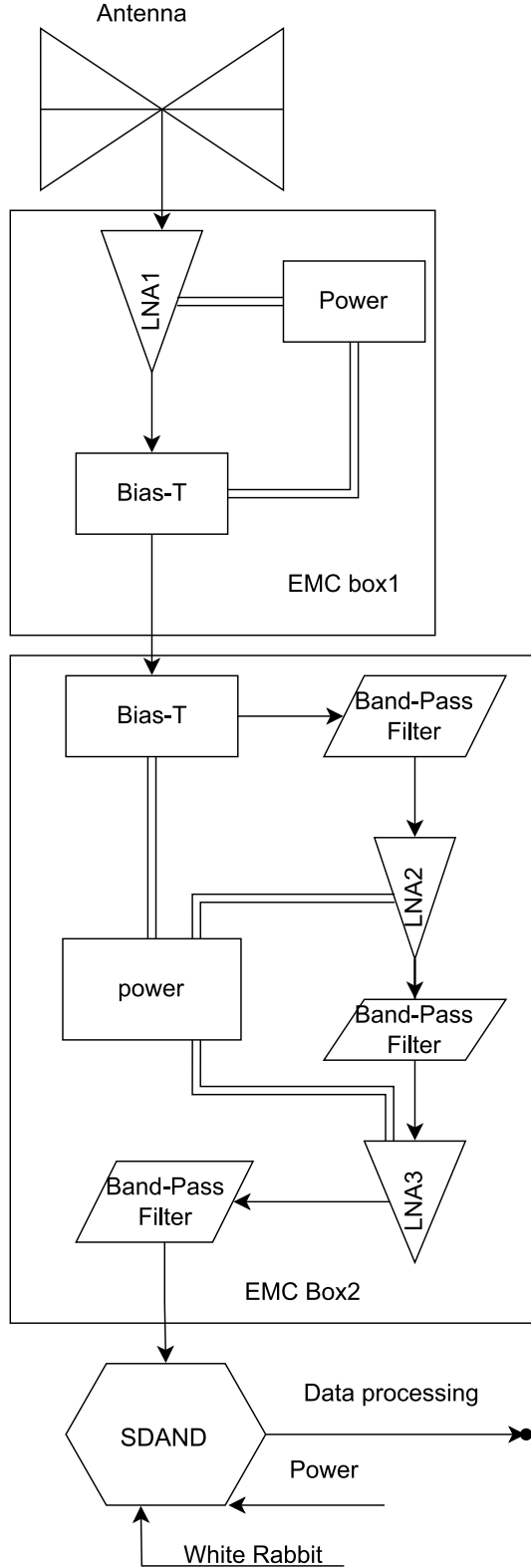


Figure 10. The block diagram showing the electronics system for one antenna in the SIGMA system.

distributed among the four GPU cards, with each GPU processing data acquired within a quarter of the time. The correlation results from the four GPU cards are gathered and summed to form the final results of each time step, which are then stored on disks. The information about the electronic design is summarized in Table 2.

4. Cross-talk

In our short-spacing interferometer array, the distance between antennas is less than half a wavelength, meaning that the interference between antennas cannot be neglected. Thekkepattu et al. (2022) discusses how receiver noise may leak into adjacent antennas, resulting in non-zero coherent receiver noise. Kwak et al. (2023) developed a model for receiver noise due to cross-coupling between different antennas. In addition, we focus on the crosstalk from sky signals and propose a strategy to remove it.

Considering there are n antennas, formed a network which has n ports. Each antenna port is connected to an LNA. Assuming all interferences are negligible, let each port's reflected wave be denoted as k_i , and let $K = (k_1, k_2, \dots, k_n)^T$. So K refers to the sky signal being reflected by the antenna. The entire coherence matrix without crosstalk effect can be expressed as:

$$\text{Coherence Matrix} = C = \begin{bmatrix} \langle k_1 k_1^* \rangle & \langle k_1 k_2^* \rangle & \dots & \langle k_1 k_n^* \rangle \\ \langle k_2 k_1^* \rangle & \langle k_2 k_2^* \rangle & \dots & \langle k_2 k_n^* \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle k_n k_1^* \rangle & \langle k_n k_2^* \rangle & \dots & \langle k_n k_n^* \rangle \end{bmatrix} \quad (3)$$

Here k_i^* denotes the complex conjugate of k_i . C can also be written as:

$$C = \langle (k_1, k_2, \dots, k_n)^T (k_1^*, k_2^*, \dots, k_n^*) \rangle = \langle K K^* \rangle, \quad (4)$$

where K^* denotes the conjugate transpose of K . In our experiment, we are only interested in baselines shorter than half a wavelength, particularly the nearest baselines. We are specifically interested in the elements on the super-diagonal and its next two diagonals.

In practical scenarios, due to the interference from the two effects previously mentioned, the signal received by the antenna is not simply K , further analysis is required. For each LNA_i , denote its forward wave as a_i and its reflected wave as b_i . Each LNA_i also has its noise source, which can be represented as an equivalent forward wave N_i and a backward wave N_i' at the LNA_i input. Also, each port has a forward wave a_i' and a reflected wave b_i' . Suppose $S = (S_{ij})$ is the scattering matrix, where S_{ij} is the scattering coefficient from antenna j to antenna i , implying that $S_{ij}^* a_j'$ represents the scattering from antenna j to antenna i .

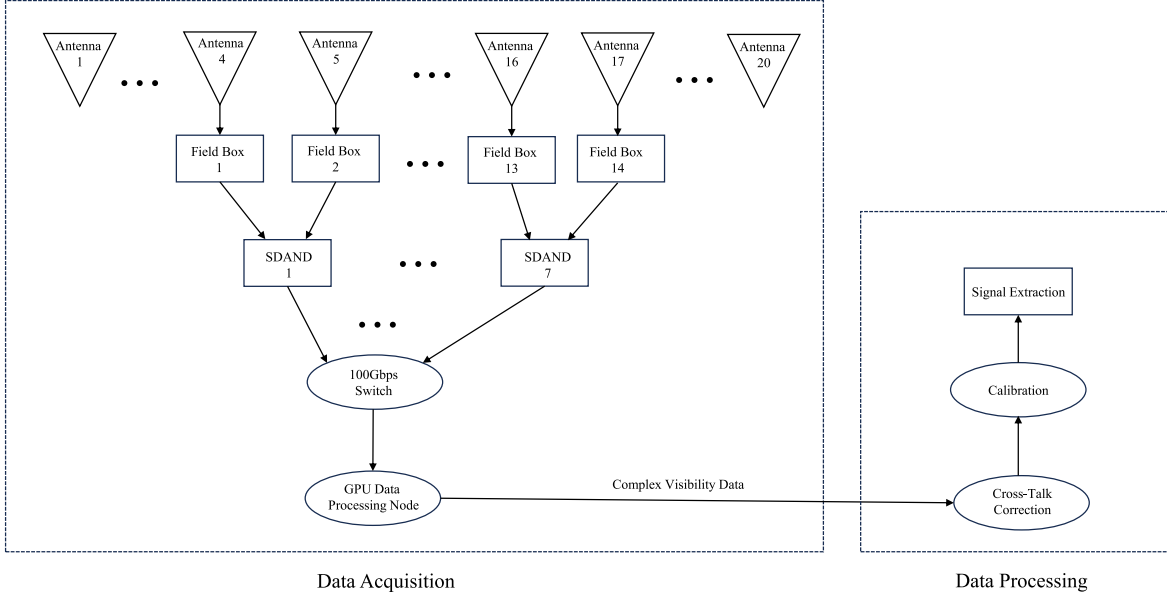


Figure 11. Block diagram of SIGMA system and data processing, antennas 4–17 will be used for data acquisition and other antennas are used as constraints, single antenna electronic system has been given in Figure 10, SDAND represents the data acquisition node, each SDAND device processes two RF inputs, which are ultimately channelized and correlated at the data processing node to get the complex visibility data for subsequent data processing.

Focusing on the two antenna ports as illustrated in Figure 12, for antenna port 1, its incident wave a_1' includes the reflected wave b_1 of LNA₁ as well as its noise's equivalent backward wave N_1' . Its reflected wave b_1' includes the original sky signal k_1 along with interference generated from antenna coupling. For LNA₁, its incident wave a_1 includes b_1' and its own noise's equivalent forward wave N_1 . It records the signal and reflects a portion of it. Let Γ_1 be the reflection coefficient of LNA₁, then the reflected signal is $b_1 = \Gamma_1 a_1$. Therefore, we have the following equations:

$$\begin{cases} a_1 = b_1' + N_1 \\ b_1 = \Gamma_1 a_1 \\ a_1' = b_1 + N_1' \\ b_1' = k_1 + S_{11}a_1' + S_{12}a_2' + \dots + S_{1n}a_n' = k_1 + \sum S_{1i}a_i' \end{cases} \quad (5)$$

We denote

$$\begin{aligned} A &= (a_1, a_2, \dots, a_n)^T, \\ B &= (b_1, b_2, \dots, b_n)^T, \\ A' &= (a_1', a_2', \dots, a_n')^T, \\ B' &= (b_1', b_2', \dots, b_n')^T, \\ N &= (N_1, N_2, \dots, N_n)^T, \\ N' &= (N_1', N_2', \dots, N_n')^T. \end{aligned}$$

Examining each port and LNA similarly, we can obtain

$$\begin{cases} A = B' + N \\ B = \Gamma A \\ A' = B + N' \\ B' = K + SA' \end{cases}, \quad (6)$$

where

$$S = \begin{pmatrix} S_{11} & S_{12} & \dots & S_{1n} \\ S_{21} & S_{22} & \dots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n1} & S_{n2} & \dots & S_{nn} \end{pmatrix}, \quad (7)$$

$$\Gamma = \begin{pmatrix} \Gamma_1 & 0 & \dots & 0 \\ 0 & \Gamma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Gamma_n \end{pmatrix}, \quad (8)$$

From (6) we can obtain

$$\begin{aligned} A &= K + SA' + N = K + SB + SN' + N \\ &= K + STA + SN' + N, \end{aligned}$$

thereby

$$(I - ST)A = K + SN' + N, \quad (9)$$

where I is the identity matrix. The matrix $I - ST$ is determined by the properties of the array itself and, as can be seen below, is assumed to be invertible in the numerical sense. Let $R = (I - ST)^{-1}$. Let $Q = SN' + N$, then Q represents the

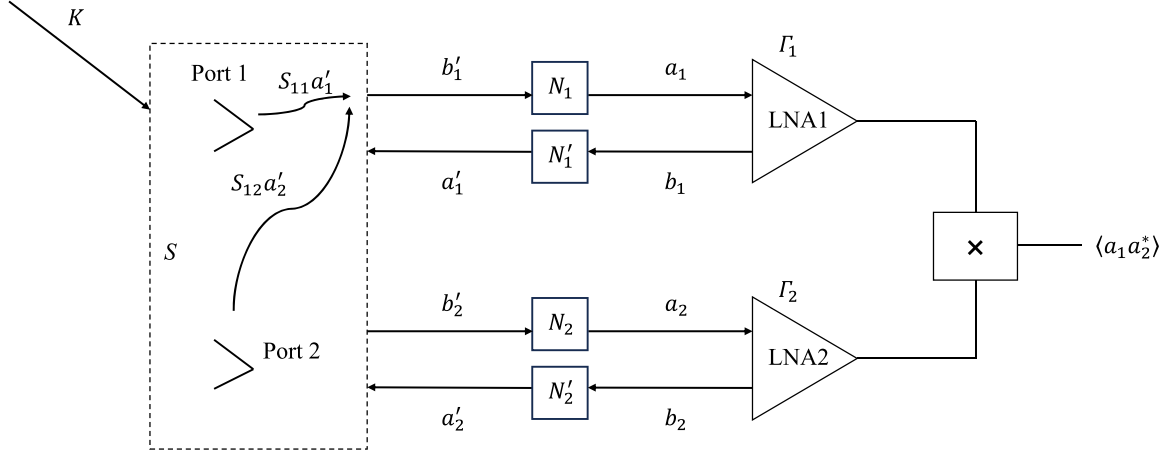


Figure 12. Simplified two-antenna interferometer model, where the signal reflects at the LNA, the reflected signal reflects again at the antenna port, and then leaks into the rest of the antenna system. For clarity, only the interference received by LNA1 is shown.

Table 2
Key Electronics Parameters and Components in the SIGMA Prototype

Component / Stage	Model / Spec	Notes
First-stage LNA	XT3074 (Chengdu Fairchild Tech)	Gain ~ 27 dB, Noise Figure (NF) ~ 0.5 dB
Second-stage amp	BFP740ESD (Infineon)	Gain ~ 40 dB, NF < 1.2 dB
Third-stage amp	ZD3027 (CEmaxRF Tech)	Gain ~ 25 dB, NF < 1.6 dB
Band-pass filter	50–200 MHz	Suppression > 50 dB outside 50–200 MHz
Number of SDAND nodes	7	Each SDAND handles 2 RF signal chains, total 14 RF inputs
Sampling scheme		0–200 MHz sampled, yields 65,536 channels (32,768 independent), ~ 6.1 kHz spectral resolution
Correlation center	GPU-based server	4 \times NVIDIA A40 GPU cards, 2 \times 100 GbE NICs
Network switch	100 GbE switch	Aggregates 7 SDAND data streams (14 total) for correlation

portion of the interference originating from the receiver noise. We have

$$A = (I - S\Gamma)^{-1}(K + SN' + N) = R(K + Q). \quad (10)$$

In practical interferometric array measurements, we are generally unable to determine A directly. Instead, we can only measure the results of their coherence.

$$\tilde{C} = \langle AA^* \rangle = R \langle (K + Q)(K^* + Q^*) \rangle R^*. \quad (11)$$

$$\begin{aligned} &= R \langle KK^* \rangle R^* + R \langle KQ^* + QK^* + QQ^* \rangle R^* \\ &= RCR^* + R \langle KQ^* + QK^* + QQ^* \rangle R^*. \end{aligned} \quad (12)$$

Our objective is to compute the undisturbed matrix C from the obtained \tilde{C} through a specific method. If we disregard the noise produced by the LNA, we can approximately obtain

$$\tilde{C} \approx RCR^*. \quad (13)$$

Evidently, to calculate C using Equation (13), it is necessary to understand the characteristics of R , which in turn implies the

need to comprehend the properties of S and Γ . Once an observational array is established, S and Γ are typically fixed and can be simulated using electromagnetic simulation software or measured in a laboratory.

Calibrating this crosstalk effect from the response of a coherent interferometer is challenging, and knowledge of the S-parameter matrix will be key to calibration. Here, we simulate the S_{nn} parameters of the SIGMA array shown in Figure 8. From Figure 8, we can see that in the operating band of our greatest interest, i.e., 65–90 MHz, the S_{nn} values for antennas 4 through 17 are nearly identical, whereas antennas 1–3 and 18–20 exhibit more noticeable deviations. In practical observations, these edge antennas (1–3 and 18–20) will not be used. Therefore, we can assume that the S-parameter matrix under this array design has class properties of the topliez matrix. The magnitude and phase plot of the simulated SIGMA S parameters matrix in 75 MHz is shown in Figure 13.

Our prior design approximated S as a Hermitian Toeplitz matrix and ensured that the diagonal elements of Γ are

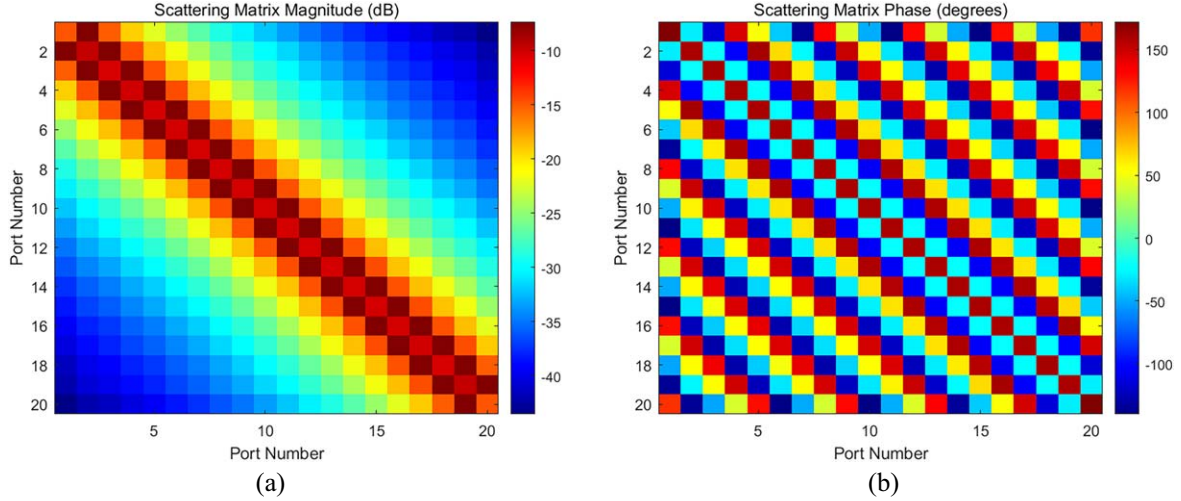


Figure 13. (a) Magnitude plot of simulated SIGMA S parameters matrix in 75 MHz. (b) Phase plot of simulated SIGMA S parameters matrix in 75 MHz. Left panel: Magnitude plot of simulated SIGMA S parameters matrix in 75 MHz. Right panel: Phase plot of simulated SIGMA S parameters matrix in 75 MHz.

approximately equal, which implies that R will also be approximately a Hermitian matrix. This implies that it can be numerically solved for its inverse or an approximate inverse, so C can be calculated by Equation (13).

5. Data Processing

Here we discuss the process of collecting data and performing processing and extraction of global signals using SIGMA, a similar process has been proposed in McKinley et al. (2020), the block diagram of SIGMA system and data processing is shown in Figure 11 and in principle there are these steps in the pipeline:

Data Acquisition. The radio-frequency signals from each antenna undergo low-noise amplification and bandpass filtering in a local field box before being passed via coaxial cables to digitization nodes. There are three low-noise amplifiers in the field box, specifically designed to minimize noise while boosting the signal strength. There is a customized 50–200 MHz bandpass filter that is applied between the amplifier stages and before analog-to-digital conversion. Once the signals are amplified and filtered, they are transmitted via coaxial cables to the digitization nodes. Each node is equipped with high-speed analog-to-digital converters (ADCs) that sample the incoming RF signals at a precise rate, converting the analog signals into digital data streams. Leveraging sub-nanosecond clock synchronization provided by the WR system, all nodes simultaneously sample the incoming signals in real-time. The digitized data streams from each node are then transferred via high-speed optical fibers to a remote GPU correlation center. This center is responsible for performing real-time channelization and correlation of the data. The channelization process involves dividing the frequency range

into multiple spectral channels, allowing for the analysis of the signal across different frequencies. The correlation process combines signals from different antennas, calculating the complex visibility at each frequency and time step. This process generates cross-correlations between pairs of antennas and auto-correlations for individual antennas. The obtained correlated complex visibility data were used for subsequent analyses.

Cross-Talk Correction. After obtaining the complex visibility, by using the known or measured system scattering parameters and the reflection coefficient of the low-noise amplifiers, we can construct a correction matrix to approximately remove the impact of crosstalk. In the absence of significant amplifier noise, the measured correlation matrix can be related to the ideal one, allowing us to invert or iteratively solve for the cleaner version of the correlation matrix. As mentioned in Section 4, once our observational array is established, its scattering matrix S , and the LNA's reflection coefficient matrix Γ , are typically fixed. S and Γ can be measured in a laboratory or be simulated using electromagnetic simulation software. Thus, we can obtain the matrix $R = (I - S\Gamma)^{-1}$. While disregarding the noise produced by the LNA, one could calculate the approximate C from our observation data \tilde{C} according to Equation (13), to reconstruct the signal of the sky.

Calibration. The complex visibility needs to be calibrated, and data from all baselines are used in the calibration process. The global sky model (De Oliveira-Costa et al. 2008) can be used as a reference in this process. The GSM sky model is a global sky model constructed from all publicly available total power large-area radio surveys, covering frequencies from 10 MHz to 100 GHz, using a principal component-based fitting

method. We can first generate GSM healpix sky maps at each frequency, and then create the reprojected map. The reprojected map is used as the model image and predicted into the MODEL column of the measurement set. Then we need to solve for the complex antenna gains using this input model. A variety of software is now available to solve antenna gain, such as CASA (Emonts et al. 2019), CALIBRATION (Offringa et al. 2016), Hyperdrive⁷ FHD (Barry et al. 2019a). As an example, the calibration in hyperdrive is based on the sky model, data visibilities are compared to the MODEL visibilities, and the differences are used to calculate antenna gains. Here is the algorithm used to determine antenna gains:

$$G_{p,i} = \frac{\sum_{q,q \neq p} D_{pq} G_{q,i-1} M_{pq}^H}{\sum_{q,q \neq p} (M_{pq} G_{q,i-1}^H) (M_{pq} G_{q,i-1}^H)^H}. \quad (14)$$

Here p and q are antenna indices, G_p is the gain Jones matrix for antenna p , D_{pq} is a data Jones matrix from baseline pq , M_{pq} is the model Jones matrix from baseline pq , i is the current iteration index, and the superscript H denotes the Hermitian transpose. This algorithm will be executed independently for each channel and once a calibration solution is obtained, the calibration solution will be applied to the input visibility data. After calibration, the data is converted to a format more suitable for further analysis.

Signal extraction after obtaining the calibrated visibility data in the appropriate format, we will select the baselines among those that are sufficiently responsive to the global signal. The length of the baseline for each pair of visibility data is obtained by calculating $\sqrt{u^2 + v^2}$, where only data shorter than 0.5 wavelengths will be used for subsequent analysis. The selected visibility data will then be converted to brightness temperature units (K) by multiplying the conversion factor $10^{-26} \lambda^2 / 2k$, where λ is the wavelength and k is the Boltzmann constant.

Each set of measured visibility fits the description of Equation (2), and the real sky contains angular structures that will affect our data extraction. We can simulate the visibility of the angular structure of the sky based on the GSM map, beam model, and baseline, and subtract it from the observed visibility, in order to mitigate the effect of the angular structure. Then we compute the global response based on the baseline and antenna radiation pattern model, and dividing it by the observed visibility gives the global sky temperature $T_{\text{sky}}(\nu)$. This process is repeated across all frequency channels to generate a global signal spectrum. In principle, after obtaining the global signal spectrum, as Bowman et al. (2018), we can use a simple polynomial foreground model to fit the data:

$$T_F(\nu) = \sum_{n=0}^{N-1} a_n \nu^{n-2.5} \quad (15)$$

where -2.5 is the exponent used to match the shape of the foreground, N is the polynomial order, a_n is the fitting coefficient, and ν is the frequency. By fitting the data in log-log space and analyzing the residuals, or jointly fitting the CD/EoR signal model with the foreground model, we can obtain an estimate of the CD/EoR signal. The method described here is the basic principle of separating angular structures and extracting CD/EoR signals from short-spacing interferometer data, simple polynomials could be unreliable enough in practical data processing (Rao et al. 2017), but can be used for technical validation.

It should be noted that further quantitative analysis and simulations are needed to verify the calibration and signal extraction procedures described here fully. The detailed estimation of calibration errors and signal extraction will be addressed in future work.

6. Discussions and Conclusions

In this paper, we present the design of the SIGMA, a novel interferometric system aimed at detecting the global 21 cm signal from the CD and the EoR. The SIGMA experiment employs a unique approach by utilizing short-spacing interferometers. In principle, a potential advantage of short-spacing interferometers is that they may be more insensitive to uncorrelated receiver noise. This work thoroughly investigates the antenna design and the SIGMA layout.

After comparing different antennas, the resulting design adopts a parallel configuration 1D array using blade antennas. We investigate the radiation pattern and the S_{nn} of the antenna array through simulations, where the mutual coupling effect when the antennas are closely spaced results in changing performance.

Another significant challenge that we face is the crosstalk between antennas. Mutual excitation between antennas is expected to introduce numerous issues in subsequent data processing, necessitating further experimentation and testing. We discuss crosstalk caused by sky signals in antenna systems, noting that these effects can be calibrated once we have a comprehensive understanding of the entire system.

The SIGMA antenna layout showcases a 1D linear array with 20 antennas, offering potential for future expansion. In the short term, our aim with SIGMA is to delve deeper into understanding the system's instrumental effects and to optimize its performance further.

Future work will focus on improving the antenna design and layout, continuing to explore mutual coupling, optimizing the receiving system, and continuing to monitor the experimental site for RFI. The feasibility of calibration and signal extraction procedures, including quantitative error estimation and foreground removal simulations, will also be investigated in future work.

⁷ https://github.com/MWATelescope/mwa_hyperdrive,

Acknowledgments

This work is supported by the National SKA Program of China (No. 2020SKA0110200 and No. 2020SKA0110100). Y.Y. acknowledges the support of the Key Program of National Natural Science Foundation of China (12433012).

Data Availability

The data underlying this article will be shared on reasonable request to the corresponding author.

ORCID iDs

Junhua Gu  <https://orcid.org/0000-0001-9765-6521>

References

- Barry, N., Beardsley, A. P., Byrne, R., et al. 2019a, *PASA*, **36**, e026
- Barry, N., Wilensky, M., Trott, C., et al. 2019b, *ApJ*, **884**, 1
- Bernardi, G., McQuinn, M., & Greenhill, L. 2015, *ApJ*, **799**, 90
- Bevins, H., Fialkov, A., de Lera Acedo, E., et al. 2022, *NatAs*, **6**, 1473
- Bowman, J. D., Rogers, A. E., & Hewitt, J. N. 2008, *ApJ*, **676**, 1
- Bowman, J. D., Rogers, A. E., Monsalve, R. A., Mozdzen, T. J., & Mahesh, N. 2018, *Natur*, **555**, 67
- de Lera Acedo, E., de Villiers, D., Razavi-Ghods, N., et al. 2022, *NatAs*, **6**, 984
- De Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., et al. 2008, *MNRAS*, **388**, 247
- Dilullo, C., Taylor, G. B., & Dowell, J. 2020, *JAI*, **9**, 2050008
- Emonts, B., Raba, R., Moellenbrock, G., et al. 2019, The CASA Software for Radio Astronomy: Status Update from ADASS 2019, arXiv:1912.09437
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *PhR*, **433**, 181
- Hothi, I., Chapman, E., Pritchard, J. R., et al. 2021, *MNRAS*, **500**, 2264
- Kolopanis, M., Pober, J. C., Jacobs, D. C., & McGraw, S. 2023, *MNRAS*, **521**, 5120
- Koopmans, L., Pritchard, J., Mellema, G., et al. 2015, in Proc. Advancing Astrophysics with the Square Kilometre Array (AASKA14) (Trieste: SISSA), 1
- Kwak, J., Podczewinski, J., Timbie, P., et al. 2023, arXiv:2311.03494
- Lipiński, M., Włostowski, T., Serrano, J., & Alvarez, P. 2011, White rabbit: a PTP application for robust sub-nanosecond synchronization, in IEEE Int. Symp. on Precision Clock Synchronization for Measurement, Control and Communication, 25 (Piscataway, NJ: IEEE)
- Martineau-Huynh, O., Bustamante, M., Carvalho, W., et al. 2017, in 7th Int. Conf. on Acoustic and Radio EeV Neutrino Detection Activities (ARENA 2016), ed. S. Buitink, J. R. Hörandel, S. de Jong et al. (Paris: EDP), 02001
- McKinley, B., Trott, C., Sokolowski, M., et al. 2020, *MNRAS*, **499**, 52
- Mertens, F. G., Mevius, M., Koopmans, L. V., et al. 2020, *MNRAS*, **493**, 1662
- Munshi, S., Mertens, F. G., Koopmans, L. V. E., et al. 2023, *A&A*, **681**, 26, arXiv:2311.05364
- Offringa, A. R., Trott, C. M., Hurley-Walker, N., et al. 2016, *MNRAS*, **458**, 1057
- Paciga, G., Albert, J. G., Bandura, K., et al. 2013, *MNRAS*, **433**, 639
- Patra, N., Subrahmanyan, R., Raghunathan, A., & Udaya Shankar, N. 2013, *ExA*, **36**, 319
- Philip, L., Abdurashidova, Z., Chiang, H., et al. 2019, *JAI*, **8**, 1950004
- Presley, M. E., Liu, A., & Parsons, A. R. 2015, *ApJ*, **809**, 18
- Rao, M. S., Subrahmanyan, R., Shankar, N. U., & Chluba, J. 2017, *ApJ*, **840**, 33
- Shaw, A. K., Chakraborty, A., Kamran, M., et al. 2023, *JApA*, **44**, 4
- Singh, S., Subrahmanyan, R., Shankar, N. U., & Raghunathan, A. 2015, *ApJ*, **815**, 88
- Sokolowski, M., Tremblay, S. E., Wayth, R. B., et al. 2015, *PASA*, **32**, e004
- Thekkepattu, J. N., McKinley, B., Trott, C. M., Jones, J., & Ung, D. C. 2022, *PASA*, **39**, e018
- Thekkepattu, J. N., Trott, C. M., & McKinley, B. 2023, *MNRAS*, **520**, 6040
- Thompson, A. R., Moran, J. M., & Swenson, G. W. 2017, *Interferometry and Synthesis in Radio Astronomy* (Berlin: Springer)
- Vedantham, H., Koopmans, L., De Bruyn, A., et al. 2015, *MNRAS*, **450**, 2291
- Voytek, T. C., Natarajan, A., Garcia, J. M. J., Peterson, J. B., López-Cruz, O., et al. 2014, *ApJL*, **782**, L9
- Zhang, X., Yue, B., Shi, Y., Wu, F., & Chen, X. 2023, *ApJ*, **945**, 109
- Zheng, Q., Wu, X.-P., Johnston-Hollitt, M., Gu, J.-h., & Xu, H. 2016, *ApJ*, **832**, 190