

<https://doi.org/10.1038/s44459-026-00068-5>

Radio resource allocation for direct-to-device services with distributed satellite systems



Giulio Orlando^{1,2}✉, Thomas Delamotte¹, Francesco Lisi^{3,4}, George Goussetis⁴, Hervé Legay³ & Andreas Knopp¹

Distributed satellite architectures are a promising solution to enhance telecommunication services over densely populated regions. However, splitting a large antenna across multiple platforms introduces significant interference challenges, particularly due to the presence of grating lobes. Previous studies have explored mitigating grating lobes by adjusting satellite positions, with encouraging results. Yet, from a resource management perspective, the optimal strategy to serve user terminals in device-to-device (D2D) applications, where high signal strength is essential for maintaining link consistency, remains unclear. This work examines a scenario where multiple platforms are arranged in a regular spatial configuration, whether physically connected in a single large satellite or distributed across several smaller ones. Leveraging the regularity of the radiated power pattern, we propose a resource allocation strategy designed to avoid interference from grating lobes. We compare the conventional N -color frequency reuse schemes with a null-grid-based approach that exploits the system's spatial regularity. Results are presented for varying numbers of colors, with additional analysis of users fairness, which does not always align with maximizing total system capacity. The findings provide valuable insights into selecting resource allocation strategies that maintain strong communication links while ensuring a more equitable distribution of service among users.

In the new space era, it is expected that billions of wireless devices will connect to LEO satellite networks, particularly through direct links between mobile phones, IoT devices, and satellites. Such connections can significantly extend coverage and improve communication efficiency, especially in remote areas or regions with limited terrestrial infrastructure. Meeting this massive connectivity demand requires the deployment of large-scale antenna arrays on satellites to enhance transmission and reception capabilities, which imposes stringent requirements on antenna performance for LEO communication systems¹. A promising architecture to address these challenges is the use of distributed satellite systems.

The concept of Distributed Satellite Systems (DSS) has been widely investigated for its ability to increase the effective antenna aperture, enhance frequency reuse, and improve performance in scenarios with highly concentrated traffic demand². The idea of satellite swarms has also been explored in ref.³, with studies showing that tethered configurations and spiral geometries can benefit from innovative deployable structures in space, initially investigated in preliminary works as early as 1999⁴. Further analyses

for both LEO and GEO have considered Formation of Arrays (FoA) as a means to improve network throughput in multibeam S-band mobile communication systems⁵. Although⁶ found FoA to be more attractive for GEO than for LEO, they nevertheless demonstrated its applicability in Cell-Free MIMO (CF-MIMO) scenarios, an aspect I leverage here to explore its potential advantages in LEO. However, in highly dynamic LEO constellations with large intersatellite distances, achieving spatial diversity gains from precoding remains challenging due to differential frequency, phase, and time offsets introduced by the distributed geometry. A more recent contribution is the Spatial Ultra-Sparse Distributed Array (SUSDA) framework⁷, which introduces a configuration design method aimed at supporting D2D services in future 6G systems. Despite these advances, all such architectures share a common challenge: the distribution of the virtual antenna aperture across multiple platforms inherently leads to grating lobe issues, which must be carefully addressed.

The literature on the grating lobe problem is extensive, with many proposed solutions aiming to suppress their occurrence despite the

¹University of the Bundeswehr Munich, Munich, Germany. ²Kongsberg NanoAvionics UAB, Vilnius, Lithuania. ³Thales Alenia Space, Toulouse, France. ⁴Heriot-Watt University, Edinburgh, UK. ✉e-mail: giulio.orlando@unibw.de

fundamental physical constraints of wave propagation. For example, a null scanning antenna (NSA) integrated into sparse phased arrays has been shown to effectively reduce sidelobe levels and scan loss while maintaining stable gain performance⁸. Similarly, grating lobe suppression in sparse radar arrays has been addressed through the reconstruction of signals from virtual antenna elements, improving angle estimation accuracy in both periodic and aperiodic configurations⁹. Another approach, the accumulated array pattern synthesis (AAPS) method combined with a measurement-based matching (MBM) algorithm, has been introduced to mitigate grating lobes in long-sparse distributed phased arrays while compensating for performance degradation caused by motion control errors¹⁰.

However, in distributed satellite architectures, especially those inspired by Formation-of-Arrays and SUSDA concepts, the inter-platform spacing is often dictated by orbital mechanics, collision-avoidance constraints, and structural considerations, and can easily exceed one wavelength at S-band. In this regime, strong grating lobes within the Field-of-View are inherently difficult to suppress due to the inherent sparsity of the large equivalent array formed by the satellite swarm. In practice, it is not always feasible to redesign the array geometry or implement advanced hardware-level suppression techniques on already-deployed systems, particularly when platform separations range from tens to hundreds of wavelengths.

In contrast to suppression-oriented methods, the present work therefore adopts a different perspective: grating lobes are treated as interference-prone forbidden regions to be explicitly avoided at the resource allocation level. Rather than attempting to eliminate grating lobes, we assume their presence and design the beam-color assignment so that users are steered away from grating-lobe directions whenever possible¹¹. Addressing this challenge fundamentally requires careful consideration of the system's resource allocation strategies, which is the focus of the proposed RRM algorithm.

Baseline allocation strategies typically assume fixed beam coverage with resources, such as power and bandwidth, divided equally among beams or user groups, without adapting to differences in traffic demand across beams. In such schemes, users may be assigned to beams without considering their spatial distribution or density, effectively clustering them arbitrarily relative to traffic hotspots or geographical correlations. More advanced baselines include classical K-means clustering. An improved K-means++ user grouping method, combined with a joint power and bandwidth allocation scheme, has been proposed for LEO satellite multicast systems, demonstrating at least a 10% capacity gain over such baseline approaches¹². However, for unicast scenarios in Distributed Satellite Systems (DSS), the literature remains extremely limited, particularly when resource allocation must contend with beam patterns and grating lobes generated by distributed arrays. Other methods involve heuristic allocation, where resources are incrementally assigned to the group or user with the highest demand or channel gain. For instance, in ref.¹³ an H-RRM scheme was introduced, where allocation decisions are made interference-wise. This approach has been applied in classical four-color frequency reuse schemes, though such schemes may not be optimal in the DSS context. Orthogonal allocation strategies (e.g. FDMA/TDMA), where users are assigned separate bandwidth or time slots, have also been explored, but these come at the cost of capacity efficiency. To address this, a multigrating allocation approach was proposed in ref.¹⁴ and ref.¹⁵, which is further expanded in the present paper to the DSS scenario. While these approaches have demonstrated significant performance gains in their respective contexts, they share a common limitation: none of them explicitly accounts for the structured interference pattern created by grating lobes. Hardware- or pattern-synthesis-based methods for grating-lobe suppression require modifications at the antenna or RF front-end level and may be difficult to deploy on resource- and mass-constrained platforms, especially once the formation geometry has been fixed. Clustering-based user grouping and classical K-means schemes optimize spatial proximity but ignore the regular structure of the radiated power pattern,

and therefore do not exploit the predictable location of grating lobes. Similarly, H-RRM was originally designed for conventional multibeam systems and does not leverage the additional degrees of freedom offered by a distributed, formation-of-arrays payload. These gaps motivate the present work, which introduce a novel coloring-based resource allocation approach designed to address the strong presence of grating lobes in DSS architectures for D2D scenarios. The proposed NG-RRM scheme is evaluated against three reference strategies: the H-RRM beam allocation method, where beams are assigned according to the half-power beamwidth (HPBW) of the radiated pattern from the Direct Radiating Array (DRA) considered in the system; a random color selection baseline, which provides a stochastic lower bound for performance; and a K-means spatial clustering baseline, which tests whether geometry-aware user grouping in the (u, v) domain can substitute interference-aware beam placement. In the DSS context, the number of colors used in the allocation is shown to significantly affect overall performance across all four strategies, and this relationship is carefully analyzed. In addition, user fairness is explicitly considered. Serving users in parallel requires not only maximizing average performance but also ensuring balanced quality of service across all users. To this end, the Jain fairness index¹⁶ is employed to monitor the resource allocation outcomes and to identify the trade-offs between throughput and fairness offered by each strategy.

The paper is structured as follows. The system model is presented in the Methods section, detailing the design of both the Space and User segments, as well as the distribution of user terminals on the ground. The Methods section also describes the grid allocation approaches for both traditional (HPBW-based) and nulls-based grids, including the multicolor strategy applied to each, introduces the KPIs selected for the analysis, and presents all four resource allocation algorithms including a formal computational complexity analysis. The numerical results and design guidelines are reported in the Results section. Finally, the findings are discussed in the Discussion section.

Notations

notation R is used as a subscript (pedix) to indicate the receiving condition, while T denotes the transmitting one. $[A]^c$ indicates a column selection of the matrix A . $[A]_a$ indicates a row selection of the matrix A . When signals are operating at a specific color frequency c , vectors or a matrices are denoted as $\mathbf{x}^c$ or A^c .

Methods

Distributed satellite system model

The scenario introduces a space segment involving a satellite in LEO. The selected satellite payload is a DRA, a phased antenna array that provides the flexibility to steer the antenna beam toward a specific direction by digitally or analogically adjusting the phase of the transmitted signals. The satellite system features a multibeam design integrated into a fully digital on-board computer, enabling multiple beams to be active simultaneously and serve different spots within the Field of View (FoV). A distribution of user terminals is generated within the FoV according to specific density rules to simulate high-density urban, rural or emergency area cores. Each UT location in the FoV is defined by spherical coordinates (θ, ϕ) , representing the angle of departure of the signal from the array's center to the user's position (see Fig. 1). The model allows us to define the type of antenna payload for the space segment, determine the number and distribution of user terminals, and select the resource allocation strategy to be implemented^{17,18}.

A LEO satellite operates at an altitude H_s . A minimum elevation angle ϵ is defined to maximize the FoV angle of aperture θ_{max} (see Eq. (1) where R_E represents the Earth's radius.), avoiding signal blockage from urban obstacles such as buildings or trees when near the horizon, accounting for Earth's curvature.

Both the altitude and θ_{max} determine the cone of line-of-sight coverage for the service area, the slant range to any point within the FoV, and the

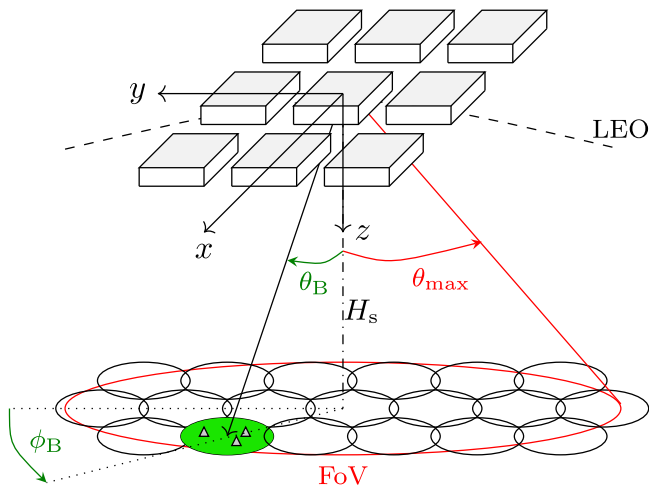


Fig. 1 | System architecture model for the distributed satellite scenario. The figure shows a LEO satellite above a spherical Earth surface. The satellite FoV is represented by the cone of solid angle delimited by the maximum scan angle θ_{max} , centered at the nadir direction. The beam pointing direction is identified by the azimuth angle ϕ_B and elevation angle θ_B , indicated on the figure. The satellite altitude H_s and the local Cartesian coordinate axes x, y, z at the satellite position are also shown.

atmospheric losses experienced by those points.

$$\theta_{max} = \arcsin \left(\frac{R_E}{R_E + H_s} \right) \cos(\epsilon) \quad (1)$$

A phased antenna array is conventionally designed with a specific number of radiating elements to enhance the main beam's peak gain, while maximizing the antenna aperture to achieve narrower, more controllable beams. Narrower beams offer two key advantages. Primarily, by serving fewer terminals per beam in a uniform user distribution, they increase the bandwidth allocated to each user. Additionally, in areas of high user density, a greater number of beams can be deployed within the same coverage area, thereby enhancing overall system capacity.

Since the satellites operate in LEO, further increasing antenna size faces practical limits due to constraints in mechanical design, thermal management, accommodation, and other engineering challenges¹⁹, there is a practical limit to antenna size. A distributed architecture addresses this constraint by splitting the antenna into multiple sub-platforms, which can be conceptualized as either tethered to the same satellite, potentially assembled in orbit, or distributed among multiple cooperating satellites. This architecture is illustrated in Fig. 2.

The full antenna has a total of N_T radiating elements. The design specifies N_p platforms arranged in a regular square grid of $\sqrt{N_p} \times \sqrt{N_p}$ of size S . The key spacing D between each platform is expressed in terms of the wavelength λ of the operating frequency. This spacing determines the occurrence of grating lobes within the FoV. Each platform itself is a phased array with n_e radiating elements, also arranged in a regular $\sqrt{n_e} \times \sqrt{n_e}$ configuration, with an intra-element spacing of $d = 0.5\lambda$. This classic element spacing prevents grating lobes at the platform level. Further analysis related to the minimum elevation angle could justify increasing this spacing up to the point where grating lobes appear only at the edge of the FoV when steering toward its opposite boundary. The variable h corresponds to the margin of the platform.

To describe the positions of all radiating elements, we introduce two Cartesian coordinate systems. The first is a global coordinate system $(\hat{x}, \hat{y})$ defined in the plane of the full distributed array, with its origin at the array center. The position of platform p in this global system is $\mathbf{r}_p = x_p \hat{x} + y_p \hat{y}$, where (x_p, y_p) are the platform coordinates. The second is a local coordinate system attached to each platform, aligned with the global axes, which is used to define the positions of the platform's radiating elements. The position of

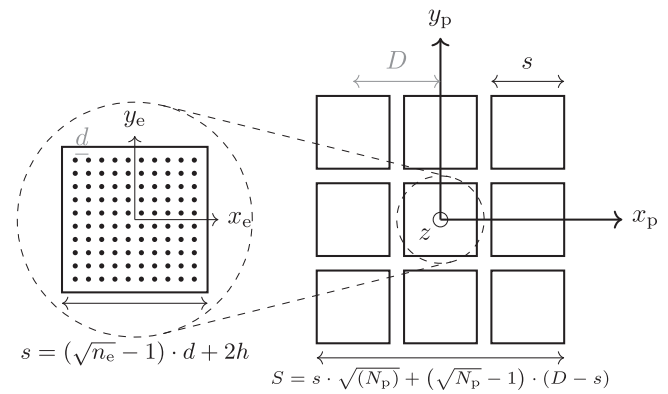


Fig. 2 | Distributed radiating array regular square lattice model. The figure illustrates the geometry of a distributed antenna system composed of N_p platforms arranged in a regular square grid of total size S , with inter-platform spacing D expressed in wavelengths. Each platform is itself a phased array of n_e radiating elements arranged in a square sub-grid with intra-element spacing $d = 0.5$ wavelengths. The platform edge length s and the array margin h are indicated. The local platform coordinate axes x_e, y_e and the global coordinate axes x, y are shown.

element e within platform p is $\mathbf{r}_e = x_e \hat{x} + y_e \hat{y}$, where (x_e, y_e) are the element coordinates relative to the platform. The absolute position of the radiating element in the global coordinate system is then given by

$$\mathbf{r} = \mathbf{r}_p + \mathbf{r}_e,$$

with $\|\mathbf{r}\|$ denoting its magnitude.

The DRA design is structured across three hierarchical levels: array, subarray, and radiating element. The elementary radiating source is modeled using a theoretical cosine element pattern, expressed as:

$$p_e(\theta) = \sqrt{\Gamma} \cdot \cos^n(\theta), \quad (2)$$

where $\Gamma = 10^{\Delta/10}$, Δ denotes the directivity (in dB) of the elementary patch and $n = \frac{\Gamma-25}{4}$. These parameters are selected such that the scanning loss does not exceed 3.5 dB when the DRA steers to the edge of the FoV. This specification is consistent with patch antennas, which are widely adopted as elementary radiators in direct radiating arrays for satellite and terrestrial systems^{20,21}, and also with commercially available (COTS) solutions developed for the 5G market.

To quantify the array response, each radiating element contributes to the generation of the total radiation pattern. The phase of each element is controlled to steer the beam in the desired direction, and beamforming techniques can be applied to minimize power in specific directions, as discussed later in this paper. These techniques apply complex weights to the array response. By defining the corresponding weights w_i for each radiating element i , the total array pattern can be expressed as:

$$p(\theta, \phi) = \mathbf{p}(\theta, \phi) \mathbf{w}, \quad (3)$$

where $\mathbf{w} = [w_1, \dots, w_{N_T}]^T$ is the column vector of beamforming weights for all radiating elements, and

$$\mathbf{p}(\theta, \phi) = [p_1(\theta, \phi), \dots, p_{N_T}(\theta, \phi)] \quad (4)$$

characterizes the complex voltage at each element, given by:

$$p_i(\theta, \phi) = p_e(\theta) e^{j\mathbf{k}(\theta, \phi) \cdot \mathbf{r}_i}, \quad (5)$$

with the wave vector defined as:

$$\mathbf{k}(\theta, \phi) = \frac{2\pi}{\lambda} \begin{bmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \end{bmatrix}, \quad (6)$$

and $\mathbf{r}_i$ representing the position vector of the i -th radiating element^{20,21}.

The grating lobes' locations can be conveniently described in the (u, v) -sine space, which represents a hemispherical projection of three-dimensional directions onto a two-dimensional plane. The transformation between spherical angles (θ, ϕ) and sine coordinates (u, v) is given by

$$u = \sin \theta \cos \phi, v = \sin \theta \sin \phi, \quad (7)$$

with the inverse mapping

$$\theta = \arcsin(\sqrt{u^2 + v^2}), \phi = \text{atan2}(v, u). \quad (8)$$

The function $\text{atan2}(v, u)$ computes the four-quadrant inverse tangent of v and u , yielding the angle ϕ between the positive u -axis and the point (u, v) . This ensures that the correct quadrant of ϕ is returned based on the signs of u and v . Note that there is no one-to-one correspondence between (θ, ϕ) and (u, v) ; for example, all pairs $(0, \phi)$ map to $(u, v) = (0, 0)$. However, each (u, v) pair uniquely identifies a direction in the upper half-space ($z \geq 0$).

For a rectangular array with inter-element spacings D_x and D_y , the grating lobe positions in sine space are given by

$$\begin{aligned} u_m &= u_0 + m \frac{\lambda}{D_x}, m \in \mathbb{Z}, \\ v_n &= v_0 + n \frac{\lambda}{D_y}, n \in \mathbb{Z}, \end{aligned} \quad (9)$$

where (u_0, v_0) denotes the main beam direction, λ is the wavelength, and (m, n) are integer indices. Thus, grating lobes appear as replicas of the main beam, displaced by integer multiples of λ/D_x and λ/D_y . When the main beam is electronically steered, the grating lobes shift accordingly while maintaining a fixed offset relative to it.

To better understand the impact of the inter-platform spacing D , we compare the array response to the reference case where $D = s$ (see Fig. 2), i.e., when the distributed architecture reduces to a continuous antenna aperture where s represents the single platform size. Figure 3 shows the resulting array patterns along the $\phi = 0$ plane for different values of D . A high maximum gain of approximately 36.4 dB is achieved with $N_T = 900$ radiating elements. As D increases, the main beam becomes narrower, but grating lobes rise in amplitude. In contrast, Fig. 4 shows that increasing the total number of elements N_T , while maintaining the same inter-element spacing and virtual antenna aperture, enhances the main-lobe gain and beam sharpness without generating additional grating lobes. However, this improvement comes at the cost of substantially higher system complexity and implementation effort²⁰.

The primary objective is to define the communication parameters that enable direct connectivity to COTS handheld devices, such as smartphones, while meeting specific data rate requirements. The user equipment (UE) characteristics are based on 3GPP TR 38.821^{22,23}, targeting handheld devices. Key parameters for system-level simulations are summarized in Table 1.

The locations of user terminals within the satellite FoV are generated in the normalized (u, v) -sine space representation. Users are placed around hubs, i.e., central cluster points representing high-demand areas.

N_{hubs} denotes the number of hubs randomly placed within the coverage disk. Each hub is associated with a cluster size, represented by the standard deviation σ . For σ close to zero, users tend to be concentrated near the hub's center, whereas for $\sigma = 1$, the users are uniformly distributed across the entire FoV. For the clustered case ($\sigma \neq 1$), each of the N_U users is randomly assigned to a hub. For hub h , the corresponding user positions are generated as

$$\mathbf{u}_i = \mathbf{u}_h + \sigma \mathbf{z}, \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2),$$

where $\mathbf{u}_h$ is the hub center. A rejection test ensures that all generated points lie within the coverage disk ($\|\mathbf{u}_i\| \leq \theta_{\text{max}}$).

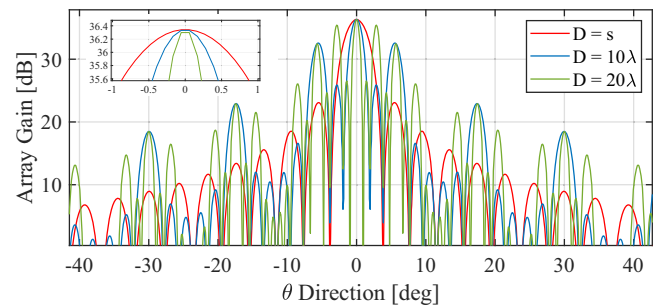


Fig. 3 | Array radiation pattern of a distributed radiating array with $N_T = 900$ elements for increasing inter-platform spacing D . The figure shows the one-dimensional cut of the array factor along the $\phi = 0$ degree plane as a function of the steering coordinate θ , for three values of inter-platform spacing: $D = s$ (solid red line), $D = 10$ (solid blue line), and $D = 20$ (solid green line), where s denotes the single-platform size and spacings are expressed in wavelengths. As D increases, the main beam becomes narrower and grating lobes of increasing amplitude appear within the field of view. The maximum gain of approximately 36.4 dBi is achieved with $N_T = 900$ elements.

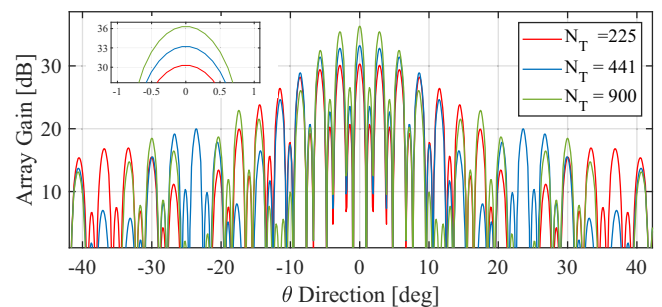


Fig. 4 | Array radiation pattern of a distributed radiating array with fixed inter-platform spacing $D = 10$ wavelengths for increasing total number of radiating elements N_T . The figure shows the one-dimensional cut of the array factor along the $\phi = 0$ degree plane as a function of the steering coordinate θ , for three values of total element count: $N_T = 225$ (solid red line), $N_T = 441$ (solid blue line), and $N_T = 900$ (solid green line). Increasing N_T sharpens the main beam and increases the peak gain without generating additional grating lobes, at the cost of higher system complexity.

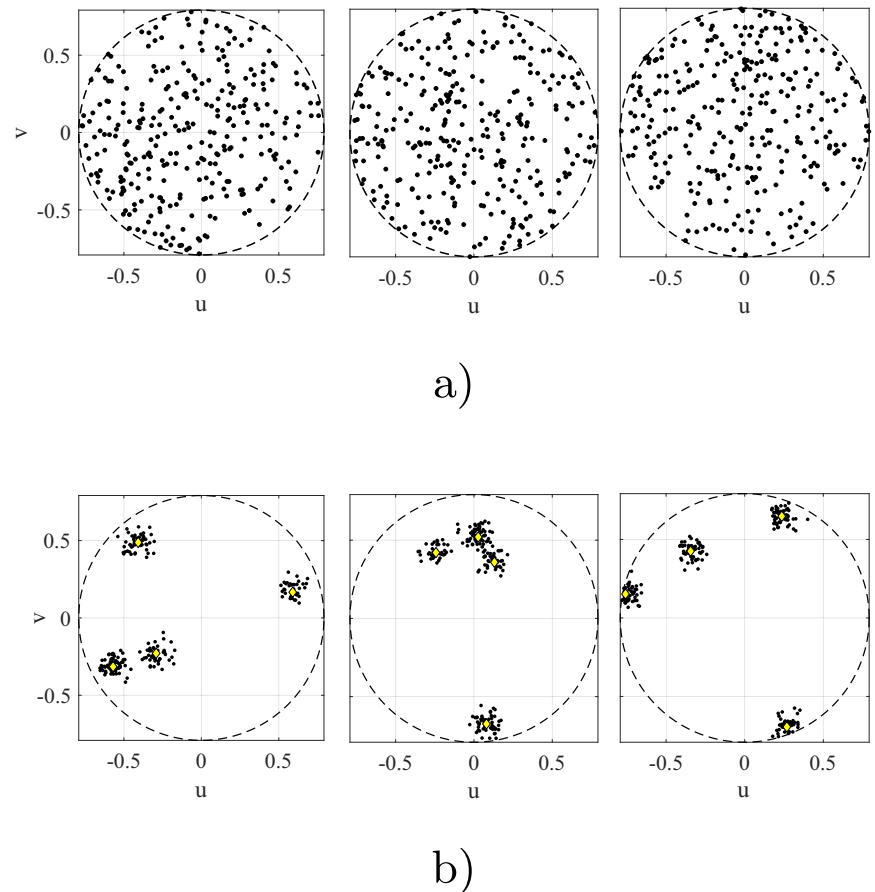
Table 1 | Handheld terminal characteristics

Characteristic	Handheld
Frequency band	S-band (2 GHz)
Antenna type	Omn-directional
Polarization	Linear ($\pm 45^\circ$ X-pol)
Rx antenna gain	0 dBi per element
Antenna temperature	290 K
Noise figure	7 dB
Tx transmit power	200 mW (23 dBm)
Tx antenna gain	0 dBi per element
G/T	-31.6 dB/K

The resulting user positions are collected in the matrix

$$\mathbf{U} = \begin{bmatrix} u_1 & v_1 \\ \vdots & \vdots \\ u_{N_U} & v_{N_U} \end{bmatrix},$$

Fig. 5 | Examples of simulated user terminal distributions in the normalized sine-space (u, v) domain, showing three independent Monte Carlo realizations each. **a Shows the uniform distribution ($\sigma = 1$), where $N_U = 300$ user terminals (represented as black dots) are distributed uniformly across the circular FoV coverage disk. **b** Shows the clustered distribution ($\sigma = 0.06$) with $N_{\text{hubs}} = 4$ hubs, where users are concentrated in four clusters. Hub centers are indicated by larger marker symbols. The circular boundary of the field of view is shown as a dashed black circle in both panels.**



containing the coordinates of all N_U users in (u, v) -space. Examples of 3 simulations are shown in Fig. 5a for the uniform distribution ($\sigma = 1$) and Fig. 5b for the clustered one ($\sigma \neq 1$).

This Gaussian cluster model is used as a first-order abstraction of urban hotspots: the number of hubs and their spreads are chosen to emulate city-like concentrations of demand embedded in a wider coverage area. While some studies^{24,25} rely on data-driven user placement based on real maps (e.g. population grids such as GPW or LandScan), here we deliberately adopt a stochastic Poisson cluster process to capture the key contrast between uniform and clustered load in a tractable way. A more detailed integration of real geographical data is left for future work; however, the main qualitative conclusions of this paper depend primarily on the presence of clustered versus uniform traffic, rather than on the exact shape of the clusters.

Grids arrangement

In a multibeam system, it is possible to predefine a grid of candidate beam directions in the (u, v) plane, where the satellite can steer beams to serve specific areas. By applying suitable precoding strategies, each beam can be designed to exhibit high SNIR regions while placing nulls in the directions of the other grid points. The digital beamforming network (D-BFN) is therefore structured as the cascade of a beam selection matrix and a fixed-beam BFN. The beam selection matrix assigns each user to one of the available fixed beams according to a specific strategy, which will be detailed later in the paper. This technique, commonly referred to as a fixed grid of beams, is widely used in GEO satellites, where the beam aperture angles are relatively small. In this study in LEO, the grid is considered at a given time instant during the satellite pass; in practice, however, the beamforming coefficients must be regularly updated to ensure that the FoV is continuously served by the same set of beams while the satellite is responsible for the considered coverage area²⁶. In this study, we adopt an adaptive

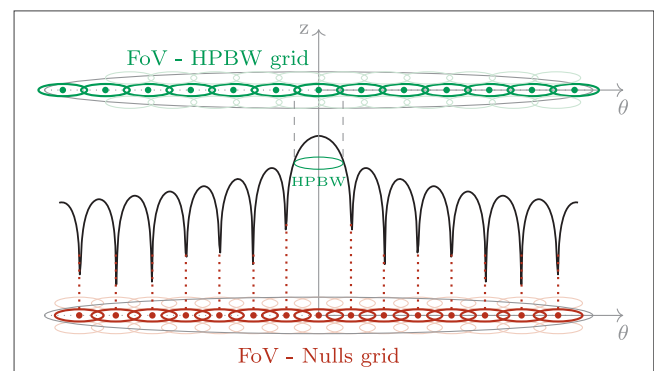


Fig. 6 | Visual concept of the two beam grid types. The figure shows a circular field of view disk with two superimposed grids of candidate beam directions. The HPBW-based grid is represented by green dots arranged in a hexagonal lattice filling the FoV. The nulls-based grid is represented by red dots arranged in a rectangular lattice, placed at the null positions of the array radiation pattern. The label z indicates the boresight direction. The label θ indicates the steering direction.

beamforming, fixed-cell perspective: beams are designed to realize a quasi-static grid of coverage cells in the (u, v) domain, and the proposed RRM algorithms operate on this fixed grid. As a result, real-time beam update rates and trajectory-dependent beam tracking, while critical in practical LEO deployments, are treated as an orthogonal design problem and are not explicitly optimized here. For a regular array structure, such as the DRA considered here, two types of grids of beams can be analysed. The methods by which these grids are defined are shown in Fig. 6 and described in the following list:

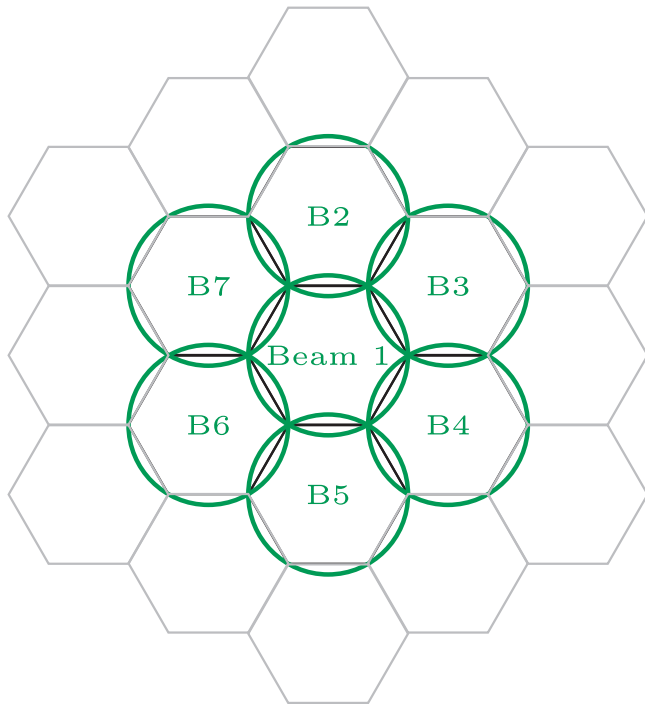


Fig. 7 | Example of a multibeam allocation based on the HPBW grid in the normalized sine-space (u, v) domain. The figure shows seven beams, labeled B1 through B7, arranged in a hexagonal lattice. Each beam is represented as a hexagonal cell. The inter-cell distance equals one HPBW unit.

- **HPBW-based Grid** - This type of grid is based on the calculation of the HPBW of the selected beam (see Fig. 6 in green). The goal is to achieve full coverage of the FoV by arranging the beams in a regular hexagonal lattice. This ensures that any user within the FoV is potentially located inside the HPBW of at least one beam. The hexagonal arrangement also introduces overlap between adjacent beams, resulting in regions covered by multiple beams. Figure 7 shows the visual concept of a hexagon disposition of beams. The distance between the centers of the beams in the hexagon results to be an HPBW unit. This can be changed in a smaller one with the purpose of having more overlapping beams inside the FoV and perform a reasonable comparison with the case of Nulls-based Grid described below.
- **Nulls-based Grid** - A “Null” refers to a point where the antenna’s radiation pattern exhibits minimal value gain. These nulls occur at specific directions relative to the antenna, where the emitted power is nearly zero or very low. In a rectangular grid arrangement, these null positions (clearly showed in Fig. 6 in red) can be used strategically to organize the placement of the beams. The pattern of nulls effectively determines possible centers where beams can be placed with minimal interference between each other at their center, displaying spots where signals from different beams are least likely to overlap and cause disruption. This is the concept of beam orthogonality and is well represented in Fig. 8. To better visualize this concept Fig. 9 shows a representation of a multibeam allocation of the nulls position in (u - v) sine space.

However, when the aperture of a DRA is fractionated, that is, divided into smaller, discrete elements, grating lobes can emerge. This arrangement can lead to these grating lobes replacing the nulls in the radiation pattern. Figure 10 illustrates two grid configurations, in the (u - v) plane, when a beam is steered to the nadir position (boresight). The top row shows the case where the element spacing D equals the platform size s , meaning there is no gap between platforms. In this configuration, the grid mapping the nulls of the radiation pattern results in a regular arrangement of points. The bottom row shows the case where $D = 20\lambda$. It is clear that some grid points are now

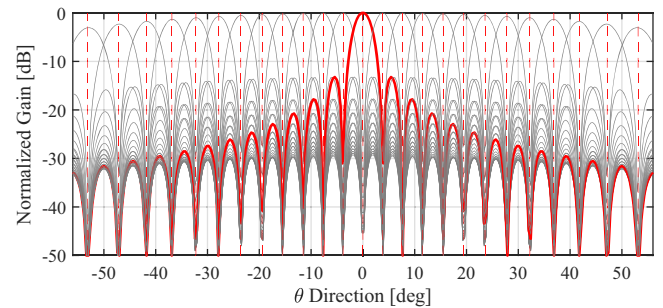


Fig. 8 | Two-dimensional cut of a multibeam allocation at the null positions for a DRA with inter-platform spacing $D = s$, illustrating the concept of beam orthogonality. The figure shows the reference beam (solid red line) steered to boresight. Null positions (dashed red lines) are marked as candidate beam center locations. Multiple beams (gray solid lines) are allocated and centered in these null positions.

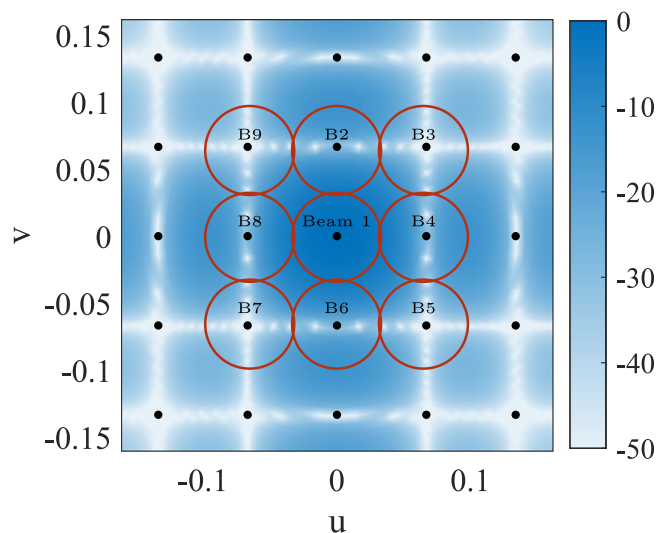


Fig. 9 | Example of a multibeam allocation at null positions in the normalized sine-space (u, v) domain. Nine beams, labeled B1 through B9, are placed at the null positions of the array radiation pattern of a DRA with $D = s$. Each beam center is shown as a labeled point on the rectangular null grid. The arrangement illustrates how the null-based grid enables spatially orthogonal beam placement across the FoV.

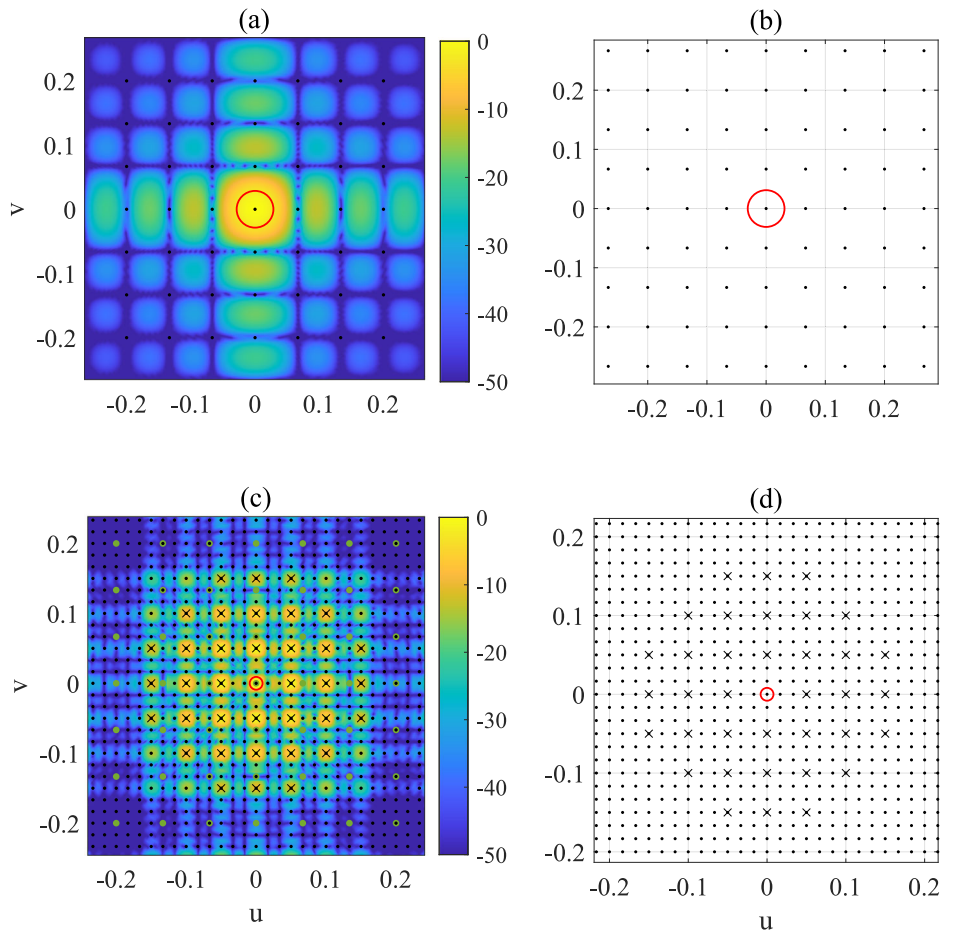
replaced by the resulting grating lobes, which are marked with an “ $\times$ ” symbol. These points can no longer be used for beam allocation.

Two main conclusions can be drawn from Fig. 10. First, increasing the distance between subarray platforms creates a denser grid of nulls within the same angular area. However, the trade-off is a significant increase in the number of grating lobes. The intensity of these grating lobes is typically stronger closer to the main beam, a characteristic governed by the radiation pattern of the radiating element, which often exhibit a cosine shape (as described in Eq. (2)).

Multicolor arrangement

Bandwidth reuse based on multibeam antennas is a key technology to achieve high capacity in High-Throughput Satellites (HTS) and to reduce the cost per bit of information delivery. The available bandwidth can be divided into sub-bands so that different sub-bands (also referred to as colors) can be allocated to different spot beams. To mitigate interference, adjacent spot beams are assigned different colors. Each color corresponds to a distinct frequency band. While the classical three- or four-color schemes are most commonly used in the literature, the coloring principle can be extended to higher-order schemes (e.g. 9, 12, 13, 16, or 19 colors).

Fig. 10 | Beam allocation grids in the (u, v) sine-space for two DRA configurations, showing the effect of inter-platform spacing on grating lobe appearance. **a, b Show the case $D = s$, where no gap exists between platforms, and all grid points correspond to null positions of the radiation pattern (represented as black dots arranged in a regular rectangular lattice). **c, d** Show the case $D = 20\lambda$, where some grid points are replaced by grating lobes, marked with a cross symbol ($\times$). These grating-lobe-contaminated positions cannot be used for beam allocation. In **(a, c)** the underlying radiation pattern is overlaid on the grid, whereas **(b, d)** remove the radiation pattern to highlight only the grid point positions on their classification (null or grating lobe).**



The methodology of applying multicolor grids differs depending on the chosen approach. For the HPBW-grid (see Fig. 11a), the grid points remain fixed, but their colors change according to the number of colors selected. Conversely, in the Nulls-grid approach (see Fig. 11b), a new grid of points is generated for each additional color introduced in the simulation. The number of sub-bands available for a given link depends on the user terminal hardware capacity as well as the user density within each spot beam. The aim of this work is to investigate the impact of increasing the number of colors in both grid configurations and to evaluate the resulting effect on system performance, with particular attention to user fairness. A comparison in the definition of the two approach follows.

The Multicolor HPBW-based Grid is based on the concept of a lattice. A lattice, $\mathcal{L}$, is a discrete set of points in an N -dimensional space, $\mathbb{R}^N$. It's defined by a set of N linearly independent basis vectors $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N\}$. These vectors can be arranged as columns in a basis matrix:

$$\mathbf{U} = [\mathbf{w}_1 | \mathbf{w}_2 | \dots | \mathbf{w}_N]$$

Any point p in the lattice $\mathcal{L}$ can be generated by taking an integer linear combination of these basis vectors. This is expressed as:

$$\mathbf{p} = \mathbf{U} \cdot \mathbf{g}$$

where $\mathbf{g}$ is a column vector of integers, $\mathbf{g} = [g_1, g_2, \dots, g_N]^T \in \mathbb{Z}^N$. The vector $\mathbf{g}$ represents the integer steps along each basis vector direction.

Considering two dimensional lattices, a specific basis matrix $\mathbf{U}$ for a 2D hexagonal grid with an inter-cell distance of δ can be defined as:

$$\mathbf{U} = \delta \cdot \begin{pmatrix} \cos(-\pi/6) & \cos(\pi/6) \\ \sin(-\pi/6) & \sin(\pi/6) \end{pmatrix} = \delta \cdot \begin{pmatrix} \sqrt{3}/2 & \sqrt{3}/2 \\ -1/2 & 1/2 \end{pmatrix}$$

Here, the two basis vectors are $\mathbf{w}_1$ and $\mathbf{w}_2$, both of length δ , separated by an angle of 60° . A value of $\delta = \sin(\theta_{3\text{dB}})/4$ has been considered in the present analysis. The smaller the value of δ , the greater the overlap of the HPBW illustrated in Fig. 7. As a consequence the fixed-grid will have more points to evaluate. The concept of “coloring” is mathematically equivalent to partitioning the main lattice $\mathcal{L}$ into a finite number of identical, interlocking sub-lattices. A sub-lattice, $\mathcal{L}'$, is a subset of points from $\mathcal{L}$ that also forms a lattice. It is defined by a new basis matrix, $\mathbf{U}'$, whose basis vectors are themselves integer linear combinations of the original basis vectors. This relationship is captured by a 2×2 integer transformation matrix, $\mathbf{M}$.

$$\mathbf{U}' = \mathbf{U} \cdot \mathbf{M}$$

where $\mathbf{M} \in \mathbb{Z}^{2 \times 2}$. The points $\mathbf{p}'$ in this primary sub-lattice (let's call it color 1'") are given by:

$$\mathbf{p}' = \mathbf{U}' \cdot \mathbf{g}' = \mathbf{U} \cdot \mathbf{M} \cdot \mathbf{g}' \text{ for } \mathbf{g}' \in \mathbb{Z}^2$$

The number of distinct sub-lattices (colors) created by this transformation is determined by the determinant of the matrix $\mathbf{M}$:

$$N_C = |\det(\mathbf{M})|$$

Geometrically, $|\det(\mathbf{M})|$ represents the ratio of the area of the new “super-cell” (the parallelogram formed by $\mathbf{U}'$) to the area of the original unit cell (formed by $\mathbf{U}$). It's the number of original lattice points contained within one super-cell. The primary sub-lattice $\mathcal{L}'$ is just one of the N_C colors. The other $N_C - 1$ colored sub-lattices are cosets of $\mathcal{L}'$. They are formed by translating $\mathcal{L}'$ by a specific set of offset vectors. Defining these offsets as

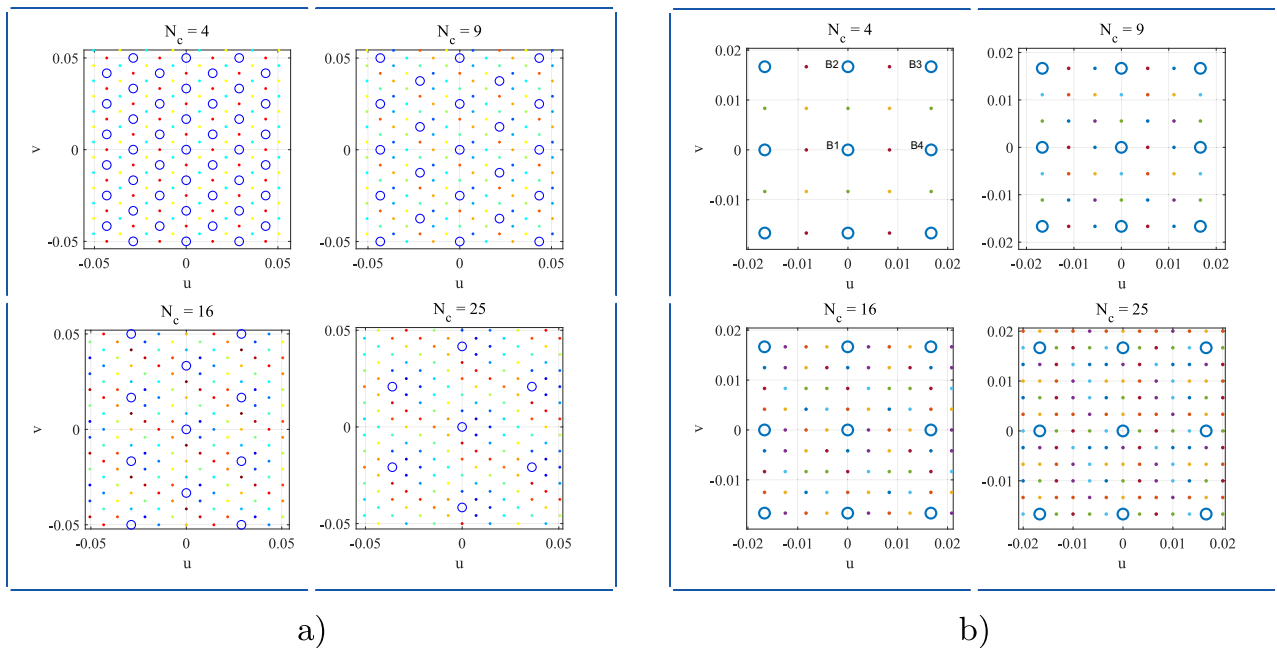


Fig. 11 | Examples of multicolor grids for the two beam placement strategies, shown for $N_c = 4, 9, 16, 25$ colors. a Shows the multicolor HBPW-based grids. Each color sub-lattice is represented by a distinct marker color, arranged in a hexagonal lattice. As N_c increases, the original lattice is partitioned into a larger number of interleaved sub-lattices. **b** Shows the multicolor nulls-based grids. Each

color grid is a rectangular sub-lattice offset from the others according to the diagonal-first ordering scheme. As N_c increases, new offset grids are added to cover the null positions more uniformly. In both panels, each color corresponds to a distinct frequency sub-band assigned to a set of non-interfering beams.

$t_c \in \mathbb{Z}^2$ for $c = 1, \dots, N_c$, the set of points for color c , denoted $\mathcal{L}_c$, is given by:

$$\mathcal{L}_c = \{\mathbf{p}' + \mathbf{U} \cdot t_c | \mathbf{p}' \in \mathcal{L}'\}$$

The set of translation vectors $\{t_c\}$ is precisely the set of integer vectors inside the parallelogram defined by the columns of $\mathbf{M}$. This ensures that every point in the original lattice $\mathcal{L}$ belongs to exactly one colored sub-lattice $\mathcal{L}_c$.

As an alternative approach the goal of Multicolor Nulls-grid method is to place the multigrids such that they cover the area (for example, defined by the four nulls B1-B4 in Fig. 9) as uniformly as possible, following the canonical color sequence. New grids are allocated starting from the diagonal positions and proceeding toward the off-diagonal points, up to the desired number of grids. Consider the task of generating a finite set of offset vectors, $\mathcal{O} = \{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_Q\}$, where Q represents the number of wanted grids. These offsets are embedded within a square region of side length γ and can be interpreted as translation vectors for positioning grid elements in a systematic manner. Let q denote the smallest integer such that $q^2 \geq Q$. This guarantees that the offsets can be placed on an $q \times q$ integer lattice,

$$\mathcal{I} = \{(i, j) | i, j \in \{0, 1, \dots, q-1\}\}.$$

Each lattice point $(i, j) \in \mathcal{I}$ corresponds to a physical coordinate.

$$\mathbf{c}_{ij} = \frac{\gamma}{q} \begin{bmatrix} i \\ j \end{bmatrix},$$

and the set of all candidate coordinates is $\mathcal{C} = \{\mathbf{c}_{ij} | (i, j) \in \mathcal{I}\}$. To impose structure, the lattice points are ordered with preference given to the main diagonal. Define the diagonal subset

$$\mathcal{I}_{\text{diag}} = \{(i, j) \in \mathcal{I} | i = j\},$$

which is sorted by increasing index i . The complementary set of off-diagonal points is $\mathcal{I}_{\text{off}} = \mathcal{I} \setminus \mathcal{I}_{\text{diag}}$. For $(i, j) \in \mathcal{I}_{\text{off}}$, the ordering is determined by the

lexicographic key

$$k(i, j) = (|i - j|, i + j, \min(i, j)).$$

The first component of the ordering key is $|i - j|$, which measures how far a point (i, j) is from the main ($i = j$). By sorting according to this value, points with smaller distance to the diagonal (i.e., those closer to $i = j$) are given priority. This strategy ensures a more uniform coverage of the area where the multigrids are placed, even when using a non-canonical number of colors. Although a more refined optimization tailored to the canonical number of colors²⁷ could further improve performance, the approach adopted here is sufficient for the purposes of this work. The complete ordered index set is obtained by concatenating diagonals and off-diagonals,

$$\mathcal{I}_{\text{ord}} = \mathcal{I}_{\text{diag}} \cup \mathcal{I}_{\text{off}},$$

and the offsets are defined as

$$\mathcal{O} = \{\mathbf{o}_k = \mathbf{c}_{i_k j_k} | (i_k, j_k) \in \mathcal{I}_{\text{ord}}, k = 1, \dots, Q\}.$$

In summary, the ordered grid offset method embeds the Q offsets in the smallest enclosing square lattice, normalizes them within a domain of side γ , and applies a diagonal-first ordering scheme. This results in a compact and structured set of translation vectors $\mathcal{O}$ suitable for systematic grid construction.

Performances modeling and KPIs

The following assumptions hold for the modeling of the system. First, the distribution of users N_U within the FoV on Earth is considered known. The transmitted symbols in $\mathbf{x}$ are assumed to be statistically independent, zero-mean, and with unit variance to ensure correct power normalization. The antennas are modeled without mutual coupling. Furthermore, it is assumed that all users experience the same noise power, and that the frequency colors employed in the system are orthogonal. To model the system of a generic

DRA that transmits the proper precoded signals $\mathbf{y}$ through N_T that radiate elements towards N_U users on the ground, the following mathematical representation is used:

$$\mathbf{z} = \sqrt{P_T} \cdot \mathbf{H}\hat{\mathbf{W}}\mathbf{x} + \mathbf{n}, \quad (10)$$

with the precoded signals defined as:

$$\mathbf{y} = \sqrt{P_T} \cdot \hat{\mathbf{W}}\mathbf{x}. \quad (11)$$

The input signals, transmitted signals, and received signals can be collected in the column vectors $\mathbf{x} = [x_1, \dots, x_j, \dots, x_{N_U}]^T$, $\mathbf{y} = [y_1, \dots, y_i, \dots, y_{N_T}]^T$ and $\mathbf{z} = [z_1, \dots, z_n, \dots, z_{N_U}]^T$, where x_j , y_i and z_i are the complex input signal for the user j , the complex signals transmitted by the radiating element l and the complex signal received by the receiver i , respectively. $\mathbf{n}$ is a vector whose elements represent the random noise process (thermal plus possible external inter-system interference) experienced by the N_U users. The channel matrix $\mathbf{H}$ of dimensions $N_U \times N_T$ describes the transfer function between the input ports of the N_T antenna radiating elements and the N_U user receivers. The generic entry element $h_{i,l}$ describes the complex transfer function between the input of the transmit radiating element l and the input of the user i receiver. The precoding matrix $\mathbf{W}$ of dimensions $N_T \times N_U$ expresses the relationship between signals at the input ports and the signals at the output ports of the precoding/beamforming On-Board Processor (OBP). The generic matrix element $w_{l,j}$ describes the complex transfer function between the input port j and output port l . The normalization of the precoding/beamforming matrix plays an important role in properly accounting for practical payload limitations such as the per antenna element power limitation affecting the DRA architecture. In the use case analysed, a CTTC normalization was applied¹³. This technique consists in a uniform per user power normalization, followed by a uniform per element power normalization. This allows to exploit all the available RF power P_T . A matching filter (MF) precoding technique, such as $\mathbf{W} = \mathbf{H}^T$, has been considered in this article. The normalized power transfer matrix $\mathbf{S}$ can be expressed as:

$$\hat{\mathbf{S}}_{u_R} = P_T \left| [\mathbf{H}]_{u_R} [\hat{\mathbf{W}}]^{u_R} \right|^2 \quad (12)$$

normalized over noise power $N = k_B T_s R_s$ with T_s the user antenna temperature, k_B the Boltzmann constant and R_s the symbol rate. Each element $s(i, j)$ of the matrix $\mathbf{S}$ quantifies the power transmitted to user j and received by user i , normalized by the thermal noise power. Diagonal entries capture the useful signal strength, such as the signal-to-noise ratio (SNR), while off-diagonal entries represent interference that needs to be minimized. The aggregate interference-to-noise ratio (INR) experienced by user i can be obtained by summing the off-diagonal elements of its corresponding row. Similarly, the off-diagonal elements of column j describe the power leaked in unintended directions by user j 's antenna pattern. In general, $\mathbf{S}$ is non-symmetric due to these directional effects.

Using the power transfer matrix, the resource allocation problem can be formulated as a color assignment task. This is naturally expressed as a 0-1 integer programming problem via a binary coloring matrix. Let N_C denote the number of available colors. The binary matrix $\mathbf{C}$, of size $N_U \times N_C$, indicates which resources are assigned to each user.

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (13)$$

The SNR, INR, and the Signal-to-Noise-plus-Interference Ratio (SNIR) for the i -th user, related to the $\mathbf{S}$ matrix normalized to noise power, can be

expressed as:

$$\text{SNR}_{u_R} = \sum_c \mathbf{C}_{u_R, c} \hat{\mathbf{S}}_{u_R, u_R}^{(c)} \quad (14)$$

$$\text{INR}_{u_R} = \sum_c \sum_{u_T \neq u_R} \mathbf{C}_{u_R, c} \mathbf{C}_{u_T, c} \hat{\mathbf{S}}_{u_R, u_T}^{(c)} \quad (15)$$

where c_R is the frequency color in reception from a user and c_T is the frequency color in transmission from the DRA.

To calculate the capacity of the system, the final SNIR referred to the receiving user u_R can be defined as:

$$\text{SNIR}_{u_R} = \frac{\text{SNR}_{u_R}}{\text{INR}_{u_R} + 1} \quad (16)$$

The Shannon capacity available for user u_R is given by:

$$C_{\text{Su}_R} = B_{u_R} \cdot \log_2 \left(1 + \text{SNIR}_{u_R} \right) \quad (17)$$

The resource allocation algorithm focuses on defining the binary coloring matrix $\mathbf{C}$ ¹³ and managing how much bandwidth B_{u_R} is available for a generic receiving user u_R , as discussed in the next chapter.

The average system throughput, T , is obtained by summing the throughput contributions of all users across each Monte Carlo realization of the user distribution within the satellite FoV. Assuming the use of the 5G-NR standard²², the throughput of the u -th user is expressed as

$$T_{u_R}^{\text{NR}} = B_{u_R} \eta_{\text{NR}} \left(\text{SNIR}_{u_R} \right), \quad (18)$$

where B_{u_R} denotes the bandwidth allocated to user u_R , and $\eta_{\text{NR}}(\cdot)$ is the spectral efficiency function defined by the 5G-NR standard, mapping the individual user signal-to-interference-plus-noise ratio SNIR_{u_R} to an achievable spectral efficiency.

Jain's Fairness index

The Jain's Fairness Index is a commonly used metric to assess the fairness of resource allocation among multiple entities. In this context, it quantifies how evenly the throughput is distributed across the satellite users¹⁶. The Jain's Fairness Index of the total throughput T in user u , with $u = 1, \dots, N_U$ is then defined as:

$$J = \frac{\left(\sum_{u=1}^{N_U} T_u \right)^2}{N_U \cdot \sum_{u=1}^{N_U} T_u^2}, \quad (19)$$

By construction, the index is bounded as follows:

$$\frac{1}{N_U} \leq J \leq 1. \quad (20)$$

A value of $J = 1$ indicates perfect fairness, meaning that the throughput is uniformly distributed across all users, whereas $J = \frac{1}{N_U}$ represents the least fair scenario, where all throughput is concentrated in a single user. In this study, the Jain's Fairness Index is evaluated for both grid configurations to determine which one performs better for a given number of color subbands.

Resource Beam-User Allocation

Table 2 summarizes the notation used throughout this section.

For Mobile Satellite Services (MSS), three specific bands are allocated as per 3GPP TS 38.101-5²⁸. These Non-Terrestrial Network satellite operating bands in FR1-NTN are summarized in Table 3.

Table 2 | Unified notation for resource allocation algorithms

Definition	Symbol
User index	$i \in \{1, \dots, N_U\}$
Color index	$c \in \{1, \dots, N_C\}$
Cluster index (K-means)	$k \in \{1, \dots, K\}, K = N_C$
Position of user i in (u, v) coordinates	$\mathbf{u}_i \in \mathbb{R}^2$
Set of grid points for color c	$\mathcal{P}^{(c)}$
Nearest grid point of user i in color c	$\mathbf{p}_i^{(c)} \in \mathcal{P}^{(c)}$
Assigned beam center for user i	$\mathbf{b}_i \in \mathbb{R}^2$
Assigned color index for user i	$a_i \in \{1, \dots, N_C\}$
Coloring matrix	$\mathbf{C} \in \{0, 1\}^{N_U \times N_C}$
State of grid point $\mathbf{p}$	$\tau(\mathbf{p}) \in \{\text{Null}, \text{Beam}, \text{Grating Lobe}\}$
Grating lobe count at grid point $\mathbf{p}$	$\gamma(\mathbf{p}) \in \mathbb{N}_0$
Grating lobe positions of a beam steered to $\mathbf{p}$	$\mathcal{G}(\mathbf{p})$
Bandwidth allocated to beam at $\mathbf{p}$	$B_{\mathbf{p}}$
Power transfer matrix	$\mathbf{S}$
Interference matrix	$\mathbf{Q} = \mathbf{S} - \text{diag}\{\text{diag}(\mathbf{S})\}$
K-means cluster centroid	$\mu_k \in \mathbb{R}^2$
K-means cluster set	$\mathcal{U}_k = \{i : a_i = k\}$
Candidate color set for user i (Random)	$\mathcal{C}_i \subseteq \{1, \dots, N_C\}$

Table 3 | Satellite downlink operating bands in FR1-NTN

NTN Satellite Operating Band	Downlink (DL) Operating Band
n256	2170–2200 MHz
n255	1525–1559 MHz
n254	2483.5–2500 MHz

The resource allocation algorithm hereafter referred to as the **Null-Grid RRM (NG-RRM)** is outlined in Algorithm 1. For each user $i \in \mathcal{U}$ and color $c \in \mathcal{C}$, let $\mathbf{p}_i^{(c)} \in \mathbb{R}^2$ denote the nearest null-grid point to user i in color c (see Fig. 12a), obtained as

$$\mathbf{p}_i^{(c)} = \arg \min_{\mathbf{q} \in \mathcal{P}^{(c)}} \|\mathbf{q} - \mathbf{u}_i\|_2. \quad (21)$$

Each grid point $\mathbf{p} \in \mathcal{P}^{(c)}$ carries a state variable $\tau(\mathbf{p}) \in \{\text{Null}, \text{Beam}, \text{Grating Lobe}\}$, initialized to Null, and a grating lobe count $\gamma(\mathbf{p})$, initialized to zero. For each user i , processed sequentially, the algorithm selects a color $c^* \in \mathcal{C}$ and assigns beam center $\mathbf{b}_i = \mathbf{p}_i^{(c^*)}$ according to the following priority rules, applied in order:

1. **Null point** ($\tau(\mathbf{p}_i^{(c)}) = \text{Null}$, see Fig. 12b): assign user i to color c , set $\tau(\mathbf{p}_i^{(c)}) \leftarrow \text{Beam}$, and for each $\mathbf{q} \in \mathcal{G}(\mathbf{p}_i^{(c)})$ set $\tau(\mathbf{q}) \leftarrow \text{Grating Lobe}$ and increment $\gamma(\mathbf{q})$ (see Fig. 10d).
2. **Existing beam** ($\tau(\mathbf{p}_i^{(c)}) = \text{Beam}$): assign user i to the existing beam at $\mathbf{p}_i^{(c)}$ without creating a new beam or updating the grating lobe map.
3. **Single grating lobe** ($\tau(\mathbf{p}_i^{(c)}) = \text{Grating Lobe}$ and $\gamma(\mathbf{p}_i^{(c)}) \leq 1$): assign user i to color c , set $\tau(\mathbf{p}_i^{(c)}) \leftarrow \text{Beam}$, update $\mathcal{G}(\mathbf{p}_i^{(c)})$ as in case 1, and halve the available bandwidth for this beam:

$$B_{\mathbf{p}_i^{(c)}} \leftarrow \frac{B_{\mathbf{p}_i^{(c)}}}{2}. \quad (22)$$

4. **Multiple grating lobes** ($\tau(\mathbf{p}_i^{(c)}) = \text{Grating Lobe}$ and $\gamma(\mathbf{p}_i^{(c)}) > 1$): remove color c from the candidate set and repeat the selection from

(21) over the reduced set $\mathcal{C} \setminus \{c\}$ until a valid assignment is found (see Fig. 12c), or all colors are exhausted. In the latter case, the closest color is assigned as a fallback.

The coloring matrix entry C_{i,c^*} is set to 1 upon assignment of user i to color c^* . The set $\mathcal{G}(\mathbf{p})$ is computed analytically from the array geometry and the beam steering direction corresponding to $\mathbf{p}$, as detailed in the Methods section. The overall complexity of NG-RRM is $\mathcal{O}(N_U N_C \log N_C)$, as derived in Table 4.

Algorithm 1. NG-RRM: Nulls-based Grid RRM

Inputs: N_C : number of colors; N_U : number of users;
 $\{\mathbf{u}_i\}_{i=1}^{N_U}$: user positions in (u, v) ; $\{\mathcal{P}^{(c)}\}_{c=1}^{N_C}$:
null-grid point sets

Outputs: $\mathbf{C} \in \{0, 1\}^{N_U \times N_C}$: coloring matrix

// Initialisation

$\tau(\mathbf{p}) \leftarrow \text{Null};$

$\gamma(\mathbf{p}) \leftarrow 0 \quad \forall \mathbf{p} \in \bigcup_c \mathcal{P}^{(c)};$

$\mathbf{C} \leftarrow \mathbf{0};$

for $i \leftarrow 1$ **to** N_U **do**

// Sort colors by distance to
nearest grid point

Compute $d_{i,c} = \|\mathbf{p}_i^{(c)} - \mathbf{u}_i\|_2$ for all c ;

$[c_1, c_2, \dots, c_{N_C}] \leftarrow \text{argsort}_c d_{i,c};$

for $c \leftarrow c_1, c_2, \dots, c_{N_C}$ **do**

if $\tau(\mathbf{p}_i^{(c)}) = \text{Null}$ **then**

$C_{i,c} \leftarrow 1;$

$\tau(\mathbf{p}_i^{(c)}) \leftarrow \text{Beam};$

$\tau(\mathbf{q}) \leftarrow \text{Grating Lobe};$

$\gamma(\mathbf{q}) += 1 \quad \forall \mathbf{q} \in \mathcal{G}(\mathbf{p}_i^{(c)});$

break;

else if $\tau(\mathbf{p}_i^{(c)}) = \text{Beam}$ **then**

$C_{i,c} \leftarrow 1;$

break;

else if $\tau(\mathbf{p}_i^{(c)}) = \text{Grating Lobe}$ ($\gamma(\mathbf{p}_i^{(c)}) \leq 1$) **then**

$C_{i,c} \leftarrow 1;$

$\tau(\mathbf{p}_i^{(c)}) \leftarrow \text{Beam};$

$\tau(\mathbf{q}) \leftarrow \text{Grating Lobe};$

$\gamma(\mathbf{q}) += 1 \quad \forall \mathbf{q} \in \mathcal{G}(\mathbf{p}_i^{(c)});$

$B_{\mathbf{p}_i^{(c)}} \leftarrow B_{\mathbf{p}_i^{(c)}}/2;$

break;

// Grating Lobe ($\gamma > 1$): try next
color in sorted list

end

if $\sum_c C_{i,c} = 0$ **then**

// Fallback: all colors
exhausted, assign closest

$C_{i,c_1} \leftarrow 1;$

end

end

The resource allocation with the HPBW-based grids follows the **H-RRM algorithm**¹³. The resource management problem is cast as a mixed-integer quadratic programming (MIQP) problem in the binary coloring

Fig. 12 | Example of user allocation using the null-grid resource and NG-RRM algorithm with $N_C = 9$ colors, illustrating the sequential beam assignment procedure of Algorithm 1. a Shows the initial grid state with all points labeled as null positions and the possible set of candidate grid points closest to the user. **b** Shows the assignment of the first user to the closest null grid point, which is then set to the beam state (filled circle). **c** Shows a case where a candidate grid point is in grating lobe state with multiple grating lobes ($\gamma > 1$), causing the algorithm to skip to the next available color. User positions are shown as colored diamonds.

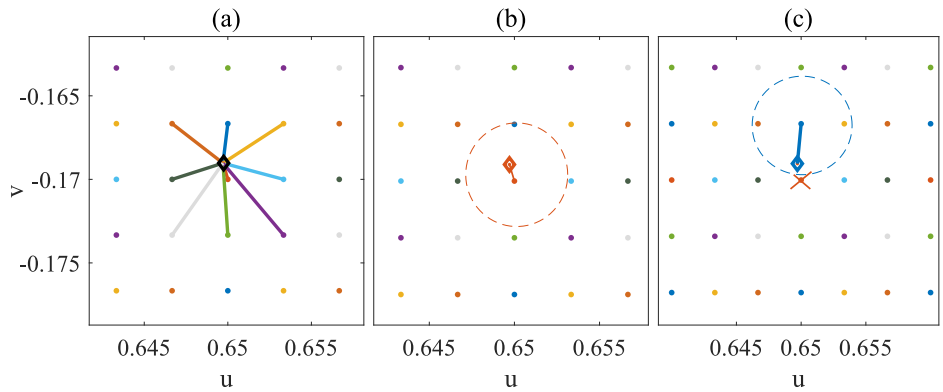


Table 4 | Computational complexity of the four allocation algorithms

Algorithm	Complexity	Notes
Random	$\mathcal{O}(N_U N_C)$	Lower bound, no spatial awareness
NG-RRM (Nulls)	$\mathcal{O}(N_U N_C \log N_C)$	Quasi-linear; GL check bounded by N_{GL}
K-means	$\mathcal{O}(l_{\max} N_U N_C)$	Linear in $l_{\max}$; $l_{\max} = 100$ in practice
H-RRM	$\mathcal{O}(N_U^2 N_C (1 + N_R))$	Quadratic in N_U ; higher fairness

matrix $\mathbf{C}$:

$$\begin{aligned} & \text{minimize } \mathbf{C} [\text{vec}(\mathbf{C})]^T (\mathbf{I} \otimes \mathbf{Q}) \text{vec}(\mathbf{C}) \\ & \text{subject to } \sum_{c=1}^{N_C} C_{i,c} = 1, C_{i,c} \in \{0, 1\}, \\ & i = 1, \dots, N_U, c = 1, \dots, N_C, \end{aligned} \quad (23)$$

where $\mathbf{C}$ is the coloring matrix defined in Table 2 and $\mathbf{Q}$ is the interference matrix:

$$\mathbf{Q} = \mathbf{S} - \text{diag}\{\text{diag}(\mathbf{S})\}, Q_{i,j} = \begin{cases} 0, & i = j, \\ S_{i,j}, & i \neq j, \end{cases} \quad (24)$$

with $\mathbf{S}$ the power transfer matrix introduced in the Methods section. Problem (23) is a MIQP in the binary variables $C_{i,c}$. Following Angeletti et al.¹³, we do not attempt to solve it via an exact MIQP solver; instead, we rely on the H-RRM heuristic, which iteratively assigns users to beams by evaluating the incremental interference cost associated with each candidate assignment. The full algorithmic details can be found in ref.¹³; here we adopt the same procedure and apply it to the HPBW-based grid considered in this work.

As an additional baseline, we consider a heuristic allocation strategy based on **K-means** clustering of users in the (u, v) domain. Let $K = N_C$ denote the number of clusters, equal to the number of available colors. The algorithm uses a nulls-based grid and proceeds in two stages. The user set $\mathcal{U}$ is partitioned into K disjoint clusters $\{\mathcal{U}_k\}_{k=1}^K$ by minimizing the within-cluster distortion

$$\min_{\{\mathcal{U}_k\}, \{\mu_k\}} \sum_{k=1}^K \sum_{i \in \mathcal{U}_k} \|\mathbf{u}_i - \mu_k\|^2, \quad (25)$$

where $\mu_k \in \mathbb{R}^2$ is the centroid of cluster $\mathcal{U}_k$. Centroids are initialized using the K-means++ strategy¹²: the first centroid is selected uniformly at random among the users, and each subsequent centroid is drawn with

probability proportional to the squared distance from the nearest already-selected centroid. The iterations then alternate between:

$$\begin{aligned} (\text{Assign}) \quad a_i &= \arg \min_{k \in \{1, \dots, K\}} \|\mathbf{u}_i - \mu_k\|^2, \\ i &= 1, \dots, N_U, \end{aligned} \quad (26)$$

$$(\text{Update}) \quad \mu_k = \frac{1}{|\mathcal{U}_k|} \sum_{i \in \mathcal{U}_k} \mathbf{u}_i \quad (27)$$

$$k = \{i : a_i = k\}, k = 1, \dots, K,$$

until the assignments $\{a_i\}$ converge. The cluster index $a_i \in \{1, \dots, K\}$ is directly mapped to a color index $c = a_i \in \{1, \dots, N_C\}$. The beam center for user i is the closest null-grid point in color a_i :

$$\mathbf{b}_i = \arg \min_{\mathbf{q} \in \mathcal{P}(a_i)} \|\mathbf{q} - \mathbf{u}_i\|_2, \quad (28)$$

and users sharing the same $\mathbf{b}_i$ and color index a_i are served by the same multibeam carrier. The coloring matrix entries are

$$C_{i,c} = \begin{cases} 1, & c = a_i, \\ 0, & \text{otherwise}, \end{cases} i = 1, \dots, N_U, c = 1, \dots, N_C. \quad (29)$$

This approach relies on purely geometric clustering for color selection, in contrast to the interference-aware optimization of (23).

As a further low-complexity baseline, each user is assigned to a color chosen uniformly at **random**, subject to a per-color load constraint. For each user i , the set of candidate colors is

$$\mathcal{C}_i = \{c \in \{1, \dots, N_C\} : N_{uc} < N_{\max}\}, N_{\max} = \left\lceil \frac{N_U}{N_C} \right\rceil + 1, \quad (30)$$

where N_{uc} is the number of users already assigned to color c . The color index is drawn as

$$a_i = \mathcal{C}_i(r), r \sim \text{Uniform}\{1, \dots, |\mathcal{C}_i|\}, i = 1, \dots, N_U, \quad (31)$$

with the convention that, if $\mathcal{C}_i$ is empty, user i is assigned to the color minimizing N_{uc} . The beam center is then the closest grid point in color a_i :

$$\mathbf{b}_i = \arg \min_{\mathbf{q} \in \mathcal{P}(a_i)} \|\mathbf{q} - \mathbf{u}_i\|_2, \quad (32)$$

and users sharing the same $(\mathbf{b}_i, a_i)$ pair are served by the same multibeam carrier. The coloring matrix entries are

$$C_{i,c} = \begin{cases} 1, & c = a_i, \\ 0, & \text{otherwise}, \end{cases} i = 1, \dots, N_U, c = 1, \dots, N_C, \quad (33)$$

providing a purely stochastic lower bound for user-color assignment performance.

Remark 1. The K-means and Random baselines are formulated here using a nulls-based grid, consistently with NG-RRM. Both can be straightforwardly adapted to the HPBW-based grid by substituting $\mathcal{P}^{(c)}$ with the corresponding HPBW grid point set.

The four allocation algorithms considered in this work differ significantly in their computational requirements. The dominant cost in each case is the user-to-color assignment step, while beam registration and grating lobe bookkeeping are shared operations of order $\mathcal{O}(N_U \cdot N_{GL})$, where N_{GL} is the number of grating lobe positions generated per beam and is bounded by a small constant determined by the array geometry.

Random allocation assigns each user to a uniformly random valid color in $\mathcal{O}(N_C)$ per user, yielding an overall complexity of $\mathcal{O}(N_U \cdot N_C)$. This represents the lowest possible complexity and serves as a computational lower bound.

NG-RRM (Nulls-grid) iterates over all N_U users and, for each user, evaluates distances to N_C color grid points and walks through the sorted candidates to find the first GL-free assignment. In the worst case all colors are inspected, giving $\mathcal{O}(N_C \log N_C)$ per user for the sorting step, plus $\mathcal{O}(N_{GL})$ per beam placement to update the grating lobe map $\gamma(\cdot)$, where $N_{GL} = |\mathcal{G}(\mathbf{p})|$ is the number of grating lobe positions within the FoV for a given beam direction. The overall complexity is therefore

$$\mathcal{O}(N_U \cdot (N_C \log N_C + N_{GL})), \quad (34)$$

where both terms contribute in practice.

K-means clustering first runs N_C initialization steps of cost $\mathcal{O}(N_U \cdot N_C)$ each (K-means++ seeding), followed by at most $I_{\max}$ iterations of assignment and centroid update, each costing $\mathcal{O}(N_U \cdot N_C)$. The subsequent beam assignment mirrors that of NG-RRM. The total complexity is therefore

$$\mathcal{O}(I_{\max} \cdot N_U \cdot N_C), \quad (35)$$

where $I_{\max} = 100$ in the implementation used here. Since $I_{\max}$ is a fixed constant, K-means scales as $\mathcal{O}(N_U \cdot N_C)$, which is of the same order as the $N_U \cdot N_C \log N_C$ term of NG-RRM for moderate N_C ; however, K-means carries a larger constant prefactor $I_{\max}$, whereas NG-RRM incurs an additional $\mathcal{O}(N_U \cdot N_{GL})$ cost for grating lobe map updates, so which algorithm is faster in practice depends on the relative magnitudes of $I_{\max}$, $\log N_C$, and N_{GL} .

H-RRM iteratively selects the user-color pair that minimises the aggregate interference cost at each step. Specifically, at each iteration the algorithm evaluates, for every unassigned user i and candidate color c , the incremental interference incurred by assigning user i to color c , computed from the interference matrix $\mathbf{Q}$ defined in (24) and the current partial coloring matrix $\mathbf{C}$. This evaluation has a per-step cost of $\mathcal{O}(N_U \cdot N_C)$. With N_R refinement steps and the main assignment loop over N_U users, the overall complexity is

$$\mathcal{O}(N_U^2 \cdot N_C \cdot (1 + N_R)), \quad (36)$$

which is quadratic in N_U . For $N_R = 0$ this reduces to $\mathcal{O}(N_U^2 \cdot N_C)$, still quadratic due to the interference-matrix update at each assignment step.

Table 4 summarizes these results. All four algorithms are polynomial-time heuristic approximations. NG-RRM offers the best trade-off between assignment quality and computational cost, scaling quasi-linearly with both N_U and N_C , while H-RRM achieves higher fairness at the cost of quadratic scaling.

Results

Performance under Uniform Distribution

Table 5 summarizes the configuration parameters of the simulations.

Table 5 | Analysis parameters

Parameter	Symbol	Value	Unit
Antenna type	–	Directive Radiating Array	–
Frequency (S-band)	f	2.17	GHz
Frequency Band	n256	30	MHz
Tx DRA element directivity	Δ	6.8	dBi
Tx DRA element spacing	d	0.5	λ
Number of radiating elements	N_T	900	–
Number of platforms	N_p	9	–
Radiating elements per platform	n_e	100	–
Platform's edge length	s	0.69	m
Array length	S	6.22	m
Array area	S^2	38.69	m ²
Platform distance	D	20	λ
Satellite orbit altitude	H_s	600	km
Max scan angle	$\theta_{\max}$	56	[°]
Total satellite output power	P_T	1260	W
Number of user terminals	N_U	100	–
Standard deviation (Uniform Distribution)	σ	1	–
Standard deviation (Clustered Distribution)	σ	0.06	–
Min number of colors	$N_{C_{\min}}$	1	#
Max number of colors	$N_{C_{\max}}$	15	#

The objective of these simulations is to analyze the behavior of resource allocation under the four considered strategies, NG-RRM, H-RRM, Random allocation, and K-means clustering, considering both uniform and clustered distributions of user terminals. The results highlight the impact of the number of color grids N_C on system throughput and fairness, and identify the most suitable algorithm for each operating scenario.

Figure 13a shows the average sum throughput as a function of N_C for the **uniform user distribution**, comparing all four allocation strategies. The four algorithms exhibit a broadly similar trend: throughput increases steeply from $N_C = 1$ to $N_C \approx 4$, driven by the progressive reduction of co-channel interference as more sub-bands become available, and then gradually saturates toward approximately 0.95 Gbps for $N_C \geq 7$. The convergence of all algorithms at high N_C reflects the fact that, when each user is effectively assigned a distinct sub-band, spatial placement strategies become equivalent and residual throughput differences are determined solely by the bandwidth fragmentation penalty. Random allocation serves as a consistent lower bound, confirming that any spatially-aware strategy outperforms unstructured assignment, while K-means and NG-RRM nearly overlap throughout the range. Figure 13b shows the Jain's fairness index for the **uniform distribution**. All four algorithms improve monotonically with N_C , converging toward $J \approx 0.90$ – 0.92 at $N_C = 10$. H-RRM is the top performer across the entire range, reaching $J \approx 0.90$ already at $N_C = 5$ and confirming its interference-aware assignment yields consistently balanced per-user throughput. Grid Nulls and Random track closely together, both achieving $J \approx 0.84$ – 0.85 at $N_C = 4$, which is a notable result: Random allocation, despite its lack of spatial intelligence, achieves fairness comparable to the deterministic Grid Nulls strategy in the uniform case, where users are evenly spread and no spatial clustering effect penalizes any particular sub-band. K-means is the clear fairness lower bound across the full range, reaching only $J \approx 0.86$ at $N_C = 10$, roughly 5–6 percentage points below H-RRM. This reveals a structural limitation of K-means in this context: by forcing users into spatially compact clusters, it tends to concentrate high-interference beams within the same color, producing uneven

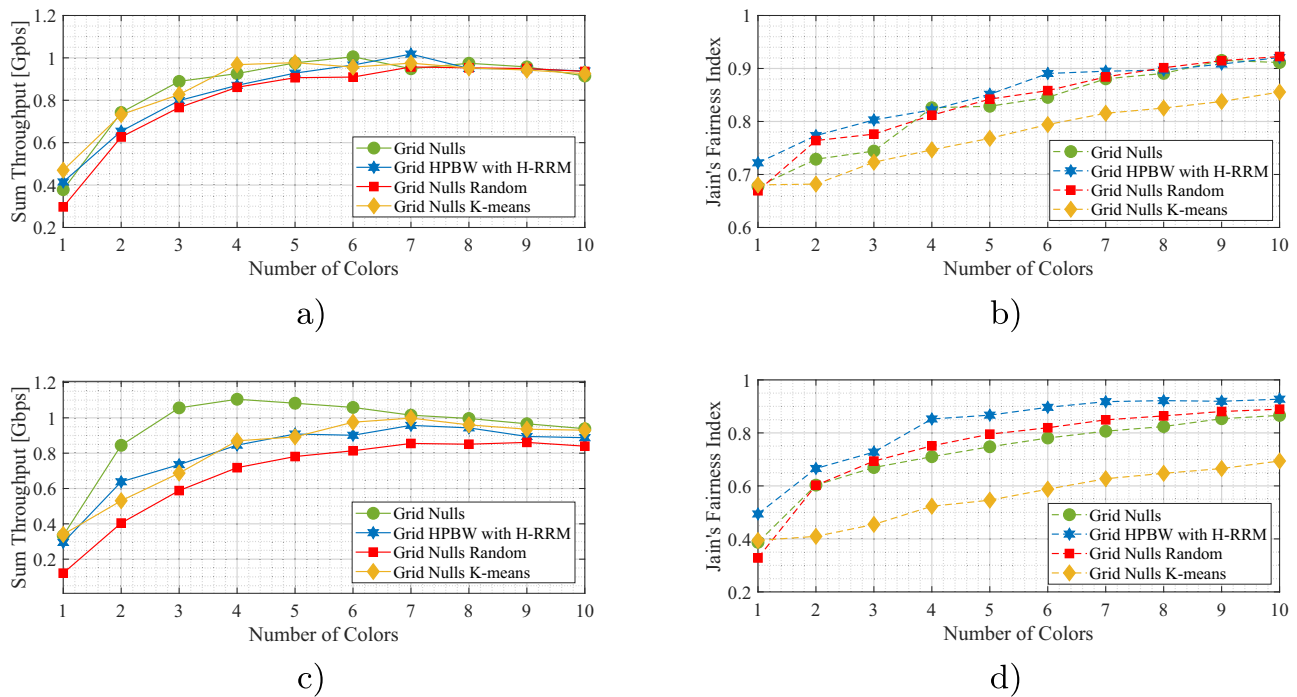


Fig. 13 | Monte Carlo simulation results for average sum throughput and Jain's fairness index as a function of the number of colors N_C , for uniform (top row) and clustered (bottom row) user distributions with $N_U = 100$ users. **a** Shows the average sum throughput in Gbps versus N_C for the uniform distribution ($\sigma = 1$), comparing four strategies: NG-RRM (solid green line with circle markers), H-RRM (solid blue line with hexagram markers), Random allocation (solid red line with square markers), and K-means clustering (solid yellow line with diamond markers).

b Shows Jain's fairness index versus N_C for the same uniform distribution and strategies, using the same line colors conventions as **a** but with dashed lines. **c** Shows the average sum throughput versus N_C for the clustered distribution ($\sigma = 0.06$), using the same four strategies and line styles. **d** Shows Jain's fairness index versus N_C for the clustered distribution. All curves are averaged over 100 independent Monte Carlo realizations of the user distribution.

per-user rates that are not corrected by the subsequent beam assignment step.

Performance under clustered distribution

Figure 13c reports the sum throughput for the **clustered user distribution**, where the algorithms diverge significantly. NG-RRM achieves the highest throughput for $N_C \leq 6$, with a clear peak at $N_C = 4$ (≈ 1.15 Gbps), representing an improvement of approximately 25–30% over H-RRM and K-means at the same operating point, and nearly $3 \times$ the throughput of Random allocation at $N_C = 1$. This behavior is directly attributable to grating-lobe-aware beam placement: in the clustered scenario, users concentrate in a small region of the FoV and grating lobes generated by beams serving one cluster tend to fall on the beam directions required for adjacent clusters. NG-RRM's explicit avoidance of such positions allows it to activate more beams with limited mutual interference at low N_C . H-RRM and K-means follow a gradual, monotonically increasing trend and converge toward NG-RRM only beyond $N_C \approx 7$, when the larger number of subbands provides sufficient spectral isolation to compensate for suboptimal spatial placement. The K-means curve closely tracks H-RRM throughout, indicating that position-based clustering in (u, v) -space does not capture the grating lobe interference structure, which is fundamentally a function of array geometry rather than user proximity alone. Figure 13d reports the Jain's fairness index for the **clustered distribution**, and reveals the most significant contrast among the four algorithms. H-RRM achieves the highest fairness across all tested configurations, reaching $J \approx 0.85$ already at $N_C = 2$ and rising smoothly to $J \approx 0.93$ – 0.95 at $N_C = 10$, confirming its robustness as a fairness-oriented strategy across all operating points. Grid Nulls and Random track each other closely in the mid range, both reaching $J \approx 0.80$ – 0.85 at $N_C = 7$ – 10 ; notably, Random allocation starts from a lower value at $N_C = 1$ ($J \approx 0.33$) but catches up with Grid Nulls by $N_C = 3$. K-means is again the fairness lower bound by a wide margin, achieving only $J \approx 0.40$ at

$N_C = 1$ and reaching $J \approx 0.70$ at $N_C = 10$, more than 20 percentage points below H-RRM. This large gap confirms that K-means clustering in the spatial domain, while effective at reducing inter-cluster co-channel interference, systematically produces unequal intra-cluster resource allocation, penalizing users at cluster boundaries whose closest beam centers coincide with grating lobe positions of nearby beams.

Taken together, the four-algorithm comparison supports the following design guidelines:

- **Throughput-critical, clustered traffic:** NG-RRM with $N_C \approx 4$ offers the best trade-off, outperforming all baselines by up to 30% in sum throughput.
- **Fairness-critical applications:** H-RRM is the preferred choice in both distributions, consistently achieving $J \geq 0.85$ from $N_C = 2$ onward and reaching near-unity fairness at moderate N_C .
- **K-means is the fairness lower bound** in both scenarios despite offering competitive throughput, exposing the structural incompatibility between spatial user clustering and grating-lobe-aware resource allocation in the DSS context.
- **Random allocation** is always the throughput lower bound but, importantly, achieves fairness comparable to NG-RRM in the uniform distribution, confirming that the throughput advantage of the null-grid strategy does not come at the expense of per-user fairness when users are evenly spread.

Furthermore, it is interesting to examine the number of beams allocated in each scenario. It is well known that the number of beams in a multibeam satellite system strongly influences the processing complexity. Increasing the number of beams affects several aspects of the system: the computation of the precoding matrix, which for linear schemes requires matrix inversion with complexity on the order of $\mathcal{O}(N_b^3)$; the

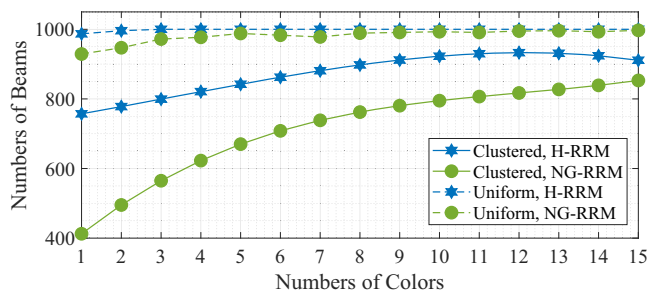


Fig. 14 | Monte Carlo simulation of the average number of active beams as a function of the number of colors N_C , comparing NG-RRM and H-RRM strategies for $N_U = 1000$ users in uniform and clustered distributions. The NG-RRM strategy under uniform distribution is shown as a dashed green line, H-RRM under uniform distribution as a dashed blue line, NG-RRM under clustered distribution as a solid green line, and H-RRM under clustered distribution as a solid blue line.

acquisition of channel state information, whose overhead scales at least linearly with the number of beams and users; and the hardware requirements, as each additional beam demands more RF chains, power amplifiers, and filtering stages. Given these well-known challenges, it is of particular interest to investigate how the number of beams affects the performance of the two primary strategies, NG-RRM and H-RRM, for which the beam allocation logic differs fundamentally; the Random and K-means baselines produce beam counts that are bounded by these two and do not add further insight to this analysis. Figure 14 illustrates the number of beams allocated by each algorithm for a scenario with $N_U = 1000$ users over increasing number of grids colors. As expected, in the uniform case both algorithms allocate nearly one beam per user, resulting in a total of approximately 1000 beams. However, in the clustered user scenario, a significant difference emerges between the two approaches. As previously discussed in the fairness analysis, H-RRM continues to prioritize allocating one beam per user, leading to a larger total number of beams. In contrast, the NG-RRM allows multiple users to share the same beam, reducing the total number of beams required. This behavior also helps explain why the fairness achieved by NG-RRM is lower than that of H-RRM. This reduction in the number of beams highlights a key advantage of the NG-RRM approach: by decreasing both computational and hardware demands, it significantly lowers system complexity while still serving the users efficiently.

Sensitivity to channel estimation errors and Doppler effects

In practice, the satellite moves at approximately 7.5 km/s at an altitude of 600 km, and channel state information (CSI) is obtained through estimation procedures subject to noise and quantization errors. Both effects introduce perturbations to the precoding weights $\mathbf{W}$, and their impact on the proposed algorithm is assessed in this subsection.

We adopt a fixed-earth-cell model in which the coverage area is partitioned into static cells on the ground, and the grid of candidate beam directions is defined so as to illuminate these cells. The distributed satellite formation flies over this fixed grid. Under the tethered-formation assumption, all platforms move as a rigid body and experience a common Doppler shift on the feeder link, which can be compensated by a single global frequency correction without per-platform phase tracking. On the user link, the fixed-cell assumption implies that only the complex gains associated with each cell evolve over time; the cell layout itself does not move, and the resource allocation is tied to geographic cells rather than to moving beams. At an orbital altitude of 600 km, the beam footprint at nadir spans several tens of kilometers, so the user displacement over a typical CSI/beam update interval on the order of one second remains small compared to the cell diameter. Even for vehicular users, the resulting position mismatch is of the order of a few hundred meters, which is negligible at S-band compared to the beamwidth. These considerations

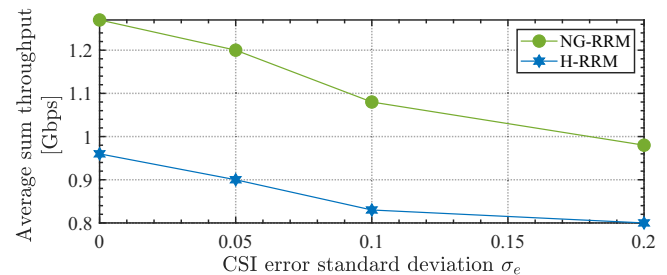


Fig. 15 | Average sum throughput in Gbps as a function of the CSI error standard deviation σ_e , for NG-RRM and H-RRM strategies under clustered user distribution with $N_C = 4$ and $N_U = 100$. The NG-RRM result is shown as a solid green line with circle markers, and the H-RRM result as a solid green line with Hexagram markers. The horizontal axis spans $\sigma_e \in \{0, 0.05, 0.1, 0.2\}$, representing increasing levels of channel estimation error normalized to the root-mean-square channel amplitude. Both algorithms exhibit graceful throughput degradation with increasing estimation error, and their relative ordering is preserved across all tested noise levels.

Table 6 | Parametric generalization: average sum throughput and Jain's fairness index for NG-RRM and H-RRM at $N_C = 4$, clustered user distribution, $N_U = 100$

Cfg	N_p	D/λ	n_e	N_T	Throughput (Gbps)		Jain's Index	
					NG-RRM	H-RRM	NG-RRM	H-RRM
A	9	20	100	900	1.20	0.97	0.73	0.91
B	16	10	49	784	1.11	0.91	0.72	0.73
C	25	50	16	400	0.29	0.23	0.65	0.86

support the use of a quasi-static channel model within each scheduling interval under a fixed-earth-cell, adaptive-beamforming framework, as also quantified in ref. ²⁶.

To further assess robustness, the estimated channel is modeled as

$$\hat{\mathbf{H}} = \mathbf{H} + \mathbf{E}, \quad (37)$$

where $\mathbf{E} \in \mathbb{C}^{N_U \times N_T}$ has i.i.d. entries $e_{ij} \sim \mathcal{CN}(0, \sigma_e^2)$, with σ_e normalized to the RMS channel amplitude so that it represents the relative estimation error. The precoding matrix $\mathbf{W}$ is computed from $\hat{\mathbf{H}}$, while the power transfer matrix $\mathbf{S}$ is always evaluated using the true channel $\mathbf{H}$. In both cases, positional perturbations and CSI errors, the resource allocation grid is unaffected. The null positions and grating lobe locations are deterministic functions of the array geometry and the beam steering direction, independent of instantaneous channel quality. Only the precoding weights degrade; the beam-color assignment remains valid.

Figure 15 reports the average sum throughput versus $\sigma_e \in \{0, 0.05, 0.1, 0.2\}$ for both the NG-RRM and H-RRM, under a clustered user distribution with $N_C = 4$ and $N_U = 100$. Both algorithms exhibit a graceful throughput degradation as estimation errors increase, and their relative ordering is preserved across all tested levels. This confirms that the throughput advantage of the proposed NG-RRM is not contingent on perfect CSI and that the algorithm is robust to moderate channel estimation errors.

Parametric generalization results

The generalization study focuses on the two primary algorithmic contributions, namely the proposed NG-RRM and the H-RRM baseline, across three alternative system configurations. The additional baselines (Random and K-means clustering) are included in the main performance comparison as lower bounds; their behavior across configurations is predictable by construction and does not require a separate parametric validation. The results in Table 6 show that the proposed NG-RRM consistently achieves higher sum throughput than H-RRM across all considered distributed array configurations, with gains on the order of 30% in the baseline case (Cfg A)

Table 7 | Design guidelines from the four-strategy comparison

Scenario	Objective	Strategy	Rationale	Complexity note
Uniform, low-mid N_C	Sum throughput	NG-RRM or H-RRM	Both strategies converge toward ≈ 0.95 Gbps for $N_C \geq 7$; difference is marginal.	NG-RRM preferred if on-board processing is constrained; H-RRM if fairness is also required.
Uniform, any N_C	Fairness	H-RRM	Highest Jain's index across the full N_C range, reaching $J \approx 0.90$ – 0.92 at $N_C = 10$.	$\mathcal{O}(N_C^2 N_C)$; quadratic scaling may be limiting for large N_U .
Clustered, $N_C \leq 6$	Sum throughput	NG-RRM with $N_C = 4$	Peak throughput ≈ 1.15 Gbps at $N_C = 4$, about 25–30% above H-RRM.	$\mathcal{O}(N_U(N_C \log N_C + N_{aL}))$; quasi-linear, suitable for resource-constrained payloads.
Clustered, any N_C	Fairness	H-RRM	Jain's index exceeds $J \approx 0.85$ from $N_C = 2$ and reaches $J \approx 0.93$ – 0.95 at higher N_C .	$\mathcal{O}(N_C^2 N_C)$; NG-RRM is preferred if complexity is a binding constraint and a moderate fairness reduction is acceptable.

and comparable improvements when increasing the number of platforms (Cfig B) or when operating with a sparser aperture and more pronounced grating lobes (Cfig C). In terms of fairness, H-RRM generally attains a higher Jain's index, reflecting its tendency to preserve a near one-beam-per-user regime whenever possible, whereas NG-RRM trades part of this fairness for higher spectral efficiency by allowing beam sharing and more aggressive use of the available colors. Overall, the relative behavior of the two algorithms, higher throughput for NG-RRM and higher fairness for H-RRM, remains consistent across different values of N_p , D/λ , and n_e , which supports the generalizability of the conclusions beyond the reference $N_p = 9$, $N_T = 900$ configuration.

Discussion

The study demonstrates that distributed satellite systems can effectively enhance direct-to-device connectivity in dense traffic areas, provided that interference caused by grating lobes is carefully managed. In this context, we compared four allocation strategies built on two grid concepts: an HPBW-based grid combined with H-RRM, a nulls-based grid combined with the proposed NG-RRM, and two baselines (Random and K-means). The numerical results show that, for uniform user distributions, all structured schemes converge to a similar throughput when a sufficiently large number of colors is available, but NG-RRM and K-means achieve slightly higher sum throughput than H-RRM in the low-moderate N_C regime, while H-RRM remains the most robust in terms of fairness. For clustered distributions, NG-RRM provides the highest sum throughput at low N_C (with gains of up to 25–30% over H-RRM and K-means around $N_C = 4$), confirming the benefit of explicitly avoiding grating-lobe directions in interference-limited, hotspot scenarios. However, H-RRM consistently achieves the best Jain's fairness index in both traffic models, highlighting a clear trade-off: NG-RRM prioritizes throughput under clustering, whereas H-RRM prioritizes fairness across the board. The analysis of the number of active beams further clarifies this trade-off. In the uniform case, both NG-RRM and H-RRM allocate close to one beam per user, leading to similar beam counts. Under clustered traffic, H-RRM continues to favor one-beam-per-user operation, increasing the total number of beams and thus the associated processing and hardware complexity. In contrast, NG-RRM naturally encourages beam sharing within clusters, reducing the number of simultaneously active beams and thereby lowering complexity at the cost of reduced fairness. The additional baselines confirm these trends: Random allocation consistently acts as a throughput lower bound (while achieving fairness close to NG-RRM in the uniform case), and K-means offers competitive throughput but systematically underperforms in fairness, exposing the limitations of purely geometric clustering in the presence of grating lobes. Taken together, these observations lead to practical design guidelines: in throughput-critical, clustered scenarios, NG-RRM with a small number of colors (e.g. $N_C \approx 4$) is preferable; in fairness-critical applications, H-RRM is the method of choice regardless of user distribution; Random and K-means serve as informative performance bounds rather than deployment candidates. These guidelines are summarized in Table 7 to assist system designers in selecting an appropriate strategy based on traffic distribution and QoS objectives. Future work should move beyond the idealized assumptions adopted here. On the channel side, it will be important to integrate more realistic user distributions derived from real datasets (e.g. population or traffic maps) and to account explicitly for LEO dynamics, in particular the rapid evolution of slant range, Doppler, and phase, which requires frequent updates of the channel matrix $\mathbf{H}$ and of the beamforming weights. On the architectural side, the relative motion of platforms, control errors, and slow variations in inter-platform spacing modify the radiated power pattern and the location and strength of grating lobes over time. Follow-on studies will therefore quantify the impact of imperfect synchronization and time-varying geometry on grating-lobe interference statistics, and investigate mitigation techniques such as adaptive grid reallocation, aperture tapering, mobility-aware precoding, and robust resource allocation, with the goal of preserving both capacity and fairness in realistic operational scenarios.

Data availability

No datasets were generated or analyzed during the current study. All numerical results reported in this work are fully reproducible from the simulation code.

Code availability

The simulation code used to generate the results of this study is available from the corresponding author on reasonable request.

Received: 2 October 2025; Accepted: 1 June 2026;

Published online: 06 August 2026

References

- Xu, Z. et al. Enhancement of satellite-to-phone link budget: An approach using distributed beamforming. *IEEE Veh. Technol. Mag.* **18**, 85–93 (2023).
- Orlando, G. et al. Optimizing communication capabilities in non-terrestrial network: A comparative analysis of continuous and distributed aperture solutions. in *Proc. ASMS/SPSC Conf.* <https://doi.org/10.5281/zenodo.14224325> (2024).
- Tuzi, D. et al. Beamforming schemes for 6G direct-to-cell connectivity using satellite swarms. in *Proc. 18th Eur. Conf. Antennas Propag. (EuCAP)*, <https://doi.org/10.23919/EuCAP60739.2024.10501021> (2024).
- Braun, A. et al. Low Earth orbit mechanical deployable structure. *Patent International Publication* WO2020/232 386 A1, 2020. [Online]. Available: [Google Patents](https://patents.google.com/patent/WO2020/232386A1/en) (2020).
- Bacci, G. et al. Formation-of-arrays antenna technology for high-throughput mobile non-terrestrial networks. *IEEE Trans. Aerosp. Electron. Syst.* **59**, no. 5 <https://doi.org/10.1109/TAES.2023.3244773> (2023).
- De Gaudenzi, R. et al. Applicability of CF-MIMO precoding to a formation of arrays (FoA) for mobile satellite communications. *IEEE Trans. Aerosp. Electron. Syst.* <https://doi.org/10.1109/TAES.2025.3556662> (2025).
- He, Y. et al. Spatial ultra-sparse array formation on LEO distributed satellite cluster: An enhanced hybrid particle swarm method. *IEEE J. Sel. Topics Signal Process.* **19**, no. 2, Mar. 2025, <https://doi.org/10.1109/JSTSP.2025.3534428>.
- Ding, Z. et al. Grating lobe suppression of sparse phased array by null scanning antenna. *IEEE Trans. Antennas Propag.* **69**, no. 11, pp. 7428–7437 <https://doi.org/10.1109/TAP.2021.3090573> (2021).
- Zhu, J. et al. Grating lobe suppression and angle estimation based on virtual antennas filling in sparse array. *IEEE Antennas Wireless Propag. Lett.* **22**, no. 3 <https://doi.org/10.1109/LAWP.2022.3216886> (2023).
- Wang, Y. et al. Grating lobe suppression for distributed phased array via accumulated array pattern synthesis. *IEEE Antennas Wireless Propag. Lett.* **22**, 7 <https://doi.org/10.1109/LAWP.2023.3249908> (2023).
- Pellet, M. et al. Uniform linear arrays with hybrid beamforming: To overlap or not to overlap? in *Proc. 19th Eur. Conf. Antennas Propag. (EuCAP)*, Stockholm, Sweden, Mar.–Apr. <https://doi.org/10.23919/EuCAP63536.2025.10999770> (2025).
- Li, Y. et al. Joint user grouping and resource allocation for LEO satellite multicast. *IEEE Syst. J.* **17**, no. 3, Sep. <https://doi.org/10.1109/JSYST.2023.3237783> (2023).
- Angeletti, P. et al. Heuristic radio resource management for massive MIMO in satellite broadband communication networks. *IEEE Access*, Oct <https://doi.org/10.1109/ACCESS.2021.3123581> (2021).
- Pellet, M. et al. Multibeam coverage optimization for broadband satellites with hybrid beamforming direct radiating arrays. in *Proc. IEEE Int. Symp. Antennas Propag.* <https://doi.org/10.1109/USNC-URSI52151.2023.10237723> (2023).
- Pellet, M. et al. Multibeam phased arrays exploiting frequency dispersion for massive MIMO satellite communications. in *Proc. 18th Eur. Conf. Antennas Propag. (EuCAP)*, Apr. <https://doi.org/10.23919/EuCAP60739.2024.10501178> (2024).
- Jain, R., Chiu, D. & Hawe, W. A quantitative measure of fairness and discrimination for resource allocation in shared systems. *Digital Equipment Corporation*, Tech. Rep. DEC-TR-301, <https://doi.org/10.48550/arXiv.cs/9809099> (1984).
- Angeletti, P. Multiple beams from planar arrays. *IEEE Trans. Antennas Propag.* **62**, no. 4 <https://doi.org/10.1109/TAP.2013.2267198> (2014).
- Angeletti, P. et al. Virtual distancing: a beam-steering technique for interference reduction in multibeam antennas. in *Proc. IEEE Int. Symp. Phased Array Syst. Technol.* <https://doi.org/10.1109/PAST49659.2022.9975078> (2022).
- AST SpaceMobile. AST SpaceMobile achieves space-based 5G cellular broadband connectivity from everyday smartphones, another historic world first. ast-science.com. Accessed: Oct. 5, 2023. [Online]. Available: <https://ast-science.com/>.
- Mailloux, R. J. *Phased Array Antenna Handbook* 2nd edn (Norwood, MA, USA: Artech House, 2005).
- Balanis, C. A. *Antenna Theory: Analysis and Design* 4th edn (Hoboken, NJ, USA: Wiley, 2016).
- 3GPP, Solutions for NR to support non-terrestrial networks (NTN), "3GPP TR 38.821, V16.2.0, Mar. 2023.
- 3GPP, Release 17, "3gpp.org. Accessed: Oct. 5, 2023. [Online]. Available: <https://www.3gpp.org/specifications-technologies/releases/release-17> 3GPP Release 17.
- Angeletti, P. et al. Optimizing massive MIMO design for non-uniform traffic in broadband telecommunication satellite networks. *IEEE Access* <https://doi.org/10.1109/ACCESS.2023.3322445> (2023).
- Vidal, F. et al. Joint precoding and resource allocation strategies applied to a large direct radiating array for GEO telecom satellite applications. in *Proc. 15th Eur. Conf. Antennas Propag. (EuCAP)* <https://doi.org/10.23919/EuCAP51087.2021.9410967> (2021).
- Momani, M. et al. Beam update rate analysis for low-complexity hybrid beamforming in LEO satellites. in *Proc. ASMS/SPSC Conf.* <https://doi.org/10.1109/ASMS/SPSC64465.2025.10946029> (2025).
- Lisi, F. et al. Interference mitigation in multibeam satellite networks as an optimal sublattice problem. *IEEE Trans. Commun.* <https://doi.org/10.1109/TCOMM.2025.3621253> (2025).
- 3GPP, Study on new radio (NR) to support non-terrestrial networks, 3GPP TR 38.811, V15.4.0, Sep. 2020.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska–Curie grant agreement No. 101072798.

Author contributions

Conceptualization: G.O., T.D. and H.L. Methodology: G.O., T.D., G.G., H.L. and A.K. Code: G.O. and F.L. Formal Analysis: T.D., G.G. and A.K. Writing: G.O. Review & Editing: All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Giulio Orlando.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026