

Cooperative Localization in Hybrid Infrared/Visible Light Networks: Theoretical Limits and Distributed Algorithms

Musa Furkan Keskin, Osman Erdem, and Sinan Gezici, *Senior Member, IEEE*

Abstract—Light emitting diode (LED) based visible light positioning (VLP) networks can provide accurate location information in environments where the global positioning system (GPS) suffers from severe signal degradation and/or cannot achieve high precision, such as indoor scenarios. In this paper, we propose to employ cooperative localization for hybrid infrared/visible light networks that involve multiple LED transmitters having known locations (e.g., on the ceiling) and visible light communication (VLC) units equipped with both LEDs and photodetectors (PDs) for the purpose of cooperation. In the considered scenario, downlink transmissions from LEDs on the ceiling to VLC units occur via visible light signals, while the infrared spectrum is utilized for device-to-device communications among VLC units. First, we derive the Cramér-Rao lower bound (CRLB) and the maximum likelihood estimator (MLE) for the localization of VLC units in the proposed cooperative scenario. To tackle the nonconvex structure of the MLE, we adopt a set-theoretic approach by formulating the problem of cooperative localization as a quasiconvex feasibility problem, where the aim is to find a point inside the intersection of convex constraint sets constructed as the sublevel sets of quasiconvex functions resulting from the Lambertian formula. Next, we devise two feasibility-seeking algorithms based on iterative gradient projections to solve the feasibility problem. Both algorithms are amenable to distributed implementation, thereby avoiding high-complexity centralized approaches. Capitalizing on the concept of quasi-Fejér convergent sequences, we carry out a formal convergence analysis to prove that the proposed algorithms converge to a solution of the feasibility problem in the consistent case. Numerical examples illustrate the improvements in localization performance achieved via cooperation among VLC units and evidence the convergence of the proposed algorithms to true VLC unit locations in both the consistent and inconsistent cases.

Index Terms— Positioning, visible light, infrared, cooperative localization, set-theoretic estimation, quasiconvex feasibility, gradient projections, quasi-Fejér convergence.

I. INTRODUCTION

A. Background and Motivation

Accurate wireless positioning plays a decisive role in various location-aware applications, including patient monitoring, inventory tracking, robotic control, and intelligent transportation systems (ITSs) [1]–[4]. In the last two decades, radio frequency (RF) based techniques have commonly been employed for wireless indoor positioning [5], [6]. Recently, light emitting diode (LED) based visible light positioning (VLP) networks have emerged as an appealing alternative to RF-based solutions, providing high-accuracy and low-cost localization services [7]. While visible light networks can be harnessed for enhancing localization performance in indoor scenarios [8], they also offer illumination and high speed data communications simultaneously without incurring additional installation costs via the use of existing LED infrastructure [9]. In VLP networks, various position-dependent parameters such as received signal

strength (RSS) [10], [11], time-of-arrival (TOA) [12], [13], time-difference-of-arrival (TDOA) [14] and angle-of-arrival (AOA) [15] can be employed for position estimation. In order to quantify performance bounds for such systems, several accuracy metrics are considered, including the Cramér-Rao lower bound (CRLB) [10], [12], [13], [16] and the Ziv-Zakai Bound (ZZB) [17].

Based on the availability of internode measurements, wireless localization networks can broadly be classified into two groups: *cooperative* and *noncooperative*. In the conventional noncooperative approach, position estimation is performed by utilizing only the measurements between anchor nodes (which have known locations) and agent nodes (the locations of which are to be estimated) [18], [19]. On the other hand, cooperative systems also incorporate the measurements among agent nodes into the localization process to achieve improved performance [19]. Benefits of cooperation among agent nodes are more pronounced specifically for sparse networks where agents cannot obtain measurements from a sufficient number of anchors for reliable positioning [20]. There exists an extensive body of research regarding the investigation of cooperation techniques and the development of efficient algorithms for cooperative localization in RF-based networks (see [18]–[20] and references therein). In terms of implementation of algorithms, centralized approaches attempt to solve the localization problem via the optimization of a global cost function at a central unit to which all measurements are delivered. Among various centralized methods, maximum likelihood (ML) and nonlinear least squares (NLS) estimators are the most widely used ones, both leading to nonconvex and difficult-to-solve optimization problems, which are usually approximated through convex relaxation approaches such as semidefinite programming (SDP) [21]–[23], second-order cone programming (SOCP) [24], [25], and convex under-estimators [26]. In distributed algorithms, computations related to position estimation are executed locally at individual nodes, thereby reinforcing scalability and robustness to data congestion [19]. Set-theoretic estimation [27]–[30], factor graphs [19], and multidimensional scaling (MDS) [31] constitute common tools employed for cooperative distributed localization in the literature.

Despite the ubiquitous use of cooperation techniques in RF-based wireless localization networks, no studies in the literature have considered the use of cooperation in VLP networks. In this study, we extend the cooperative paradigm to infrared and visible light domains. More specifically, we set forth a cooperative localization framework for hybrid infrared/visible light networks whereby LED transmitters on the ceiling function as anchors with known locations and visible light communication (VLC) units with unknown locations are equipped with LEDs and photodetectors (PDs) for the purpose of communications with both LEDs on the ceiling and other VLC units¹. Utilization

The authors are with the Department of Electrical and Electronics Engineering, Bilkent University, 06800, Ankara, Turkey, Tel: +90-312-290-3139, Fax: +90-312-266-4192, Emails: {keskin,oerdem,gezici}@ee.bilkent.edu.tr.

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the authors. The material includes the proofs and derivations. Contact gezici@ee.bilkent.edu.tr for further questions about this work.

Part of this work was presented at IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), Istanbul, Turkey, June 2017.

¹In this paper, “VLP network” (when used in the context of cooperation) and “hybrid infrared/visible light network” are used interchangeably to refer to an indoor VLC system where LED transmitters on the ceiling emit visible light signals to VLC units for downlink communications and VLC units communicate with one another using infrared wavelengths.

of the proposed framework is primarily motivated by indoor VLC systems endowed with infrared uplink capability. Since infrared LEDs and PDs are already available in some VLC systems for efficient uplink transmission [8], [32]–[35], they can also be utilized for device-to-device communications to achieve cooperation among VLC units. An additional benefit of using infrared wavelengths for cooperation is that it helps mitigate eye safety risks incurred by communications among VLC units [36]².

The proposed network facilitates the definition of arbitrary connectivity sets between the LEDs on the ceiling and the VLC units, and also among the VLC units, which can provide significant performance enhancements over the traditional noncooperative approach employed in the VLP literature. Based on the noncooperative (i.e., between LEDs on the ceiling and VLC units) and cooperative (i.e., among VLC units) RSS measurements, we first derive the CRLB and the MLE for the localization of VLC units. Since the MLE poses a challenging nonconvex optimization problem, we follow a set-theoretic estimation approach and formulate the problem of cooperative localization as a quasiconvex *feasibility* problem (QFP) [37], where feasible constraint sets correspond to sublevel sets of certain type of quasiconvex functions. The quasiconvexity arising in the problem formulation stems from the Lambertian formula, which characterizes the attenuation level of infrared/visible light channels³. Next, we design two feasibility-seeking algorithms, having cyclic and simultaneous characteristics, which employ iterative gradient projections onto the specified constraint sets. From the viewpoint of implementation, the proposed algorithms can be implemented in a distributed architecture that relies on computations at individual VLC units and a broadcasting mechanism to update position estimates. Moreover, we provide a formal convergence proof for the projection-based algorithms based on quasi-Féjer convergence, which enjoys decent properties to support theoretical analysis [38].

B. Literature Survey on Set-Theoretic Estimation

The applications of convex feasibility problems (CFPs) encompass a wide variety of disciplines, such as wireless localization [27]–[30], [39], compressed sensing [40], image recovery [41], image denoising [42] and intensity-modulated radiation therapy [43]. In contrary to optimization problems where the aim is to minimize the objective function while satisfying the constraints, feasibility problems seek to find a point that satisfies the constraints in the absence of an objective function [40]. Hence, the goal of a CFP is to identify a point inside the intersection of a collection of closed convex sets in a Euclidean (or, in general, Hilbert) space. In feasibility problems, a commonly pursued approach is to perform projections onto the individual constraint sets in a sequential manner, rather than projecting onto their intersection due to analytical intractability [44]. The work in [27] formulates the problem of acoustic source localization as a CFP and employs the well-known projections onto convex sets (POCS) technique for convergence to true source locations. Following a similar methodology, the noncooperative wireless positioning problem with noisy range measurements is modeled as a CFP in [28], where POCS and outer-approximation (OA) methods are utilized to derive distributed algorithms that perform well under non-line-of-sight (NLOS) conditions. In [39], a cooperative localization approach based on projections onto nonconvex boundary sets is proposed

²Please see Remark 2 in Section II-B for an eye-safety analysis of infrared LEDs.

³Throughout the paper, the near-infrared region between 780 nm and 950 nm is considered as the infrared spectrum [36]. The Lambertian channel attenuation model is valid in both the visible light spectrum and the infrared region from 780 nm to 950 nm [36].

for sensor networks, and it is shown that the proposed strategy can achieve better localization performance than the centralized SDP and the distributed MDS although it may get trapped into local minima due to nonconvexity. Similarly, the work in [29] designs a POCS-based distributed positioning algorithm for cooperative networks with a convergence guarantee regardless of the consistency of the formulated CFP, i.e., whether the intersection is nonempty or not.

Although CFPs have attracted a great deal of interest in the literature, QFPs have been investigated only rarely. QFPs represent generalized versions of CFPs in that the constraint sets are constructed from the lower level sets of quasiconvex functions in QFPs whereas such functions are convex in CFPs [37]. The study in [37] explores the convergence properties of subgradient projections based iterative algorithms utilized for the solution of QFPs. It is demonstrated that the iterations converge to a solution of the QFP if the quasiconvex functions satisfy Hölder conditions and the QFP is consistent, i.e., the intersection is nonempty. In this work, we show that the Lambertian model based (originally non-quasiconvex) functions can be approximated by appropriate quasiconvex lower bounds, which convexifies the (originally nonconvex) sublevel constraint sets, thus transforming the formulated feasibility problem into a QFP.

C. Contributions

The previous work on VLP networks has addressed the problem of position estimation based mainly on the ML estimator [13], [16], [45], the least squares estimator [15], [45], triangulation [11], [46], and trilateration [14] methods. This paper, however, considers the problem of localization in VLP networks as a feasibility problem and introduces efficient iterative algorithms with convergence guarantees in the consistent case. In addition, the theoretical bounds derived for position estimation are significantly different from those in [16], [45] via the incorporation of terms related to cooperation, which allows for the evaluation of the effects of cooperation on the localization performance in any three dimensional cooperative VLP scenario. Furthermore, unlike the previous research on localization in RF-based wireless networks via CFP modeling [27]–[29], [39], where a common approach is to employ POCS-based iterative algorithms, we formulate the localization problem as a QFP for VLP networks, which necessitates the development of more sophisticated algorithms (e.g., gradient projections) and different techniques for studying the convergence properties of those algorithms (e.g., quasiconvexity and quasi-Féjer convergence). The main contributions of this paper can be summarized as follows:

- For the first time in the literature, we propose to employ cooperative localization for VLP networks via a generic configuration that allows for an arbitrary construction of connectivity sets and transmitter/receiver orientations.
- The CRLB for localization of VLC units is derived in the presence of cooperative measurements (Section II). The effects of cooperation on the performance of localization in VLP networks are illustrated based on the provided CRLB expression (Section VI-A).
- The problem of cooperative localization in VLP networks is formulated as a quasiconvex feasibility problem, which circumvents the complexity of the nonconvex ML estimator and facilitates efficient feasibility-seeking algorithms (Section III).
- We design gradient projections based low-complexity iterative algorithms to find solutions to the feasibility problem (Section IV). The proposed set-theoretic framework favors the implementation of algorithms in a distributed architecture.

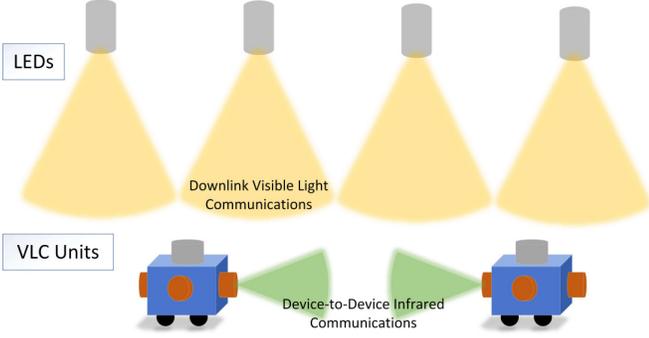


Fig. 1. Cooperative VLP network.

- We provide formal convergence proofs for the proposed algorithms in the consistent case based on the concept of quasi-Fejér convergence (Section V).

II. SYSTEM MODEL AND THEORETICAL BOUNDS

A. System Model

The proposed cooperative VLP network consists of N_L LED transmitters and N_V VLC units, as illustrated in Fig. 1. The location of the j th LED transmitter is denoted by \mathbf{y}_j and its orientation vector is given by $\tilde{\mathbf{n}}_{T,j}$ for $j \in \{1, \dots, N_L\}$. The locations and the orientations of the LED transmitters are assumed to be known, which is a reasonable assumption for practical systems [45], [46]. In the proposed system, each VLC unit not only gathers signals from the LED transmitters but also communicates with other VLC units in the system for cooperation purposes. To that aim, the VLC units are equipped with both LEDs and PDs; namely, there exist L_i LEDs and K_i PDs at the i th VLC unit for $i \in \{1, \dots, N_V\}$. The unknown location of the i th VLC unit is denoted by \mathbf{x}_i , where $i \in \{1, \dots, N_V\}$. For the j th PD at the i th VLC unit, the location is given by $\mathbf{x}_i + \mathbf{a}_{i,j}$ and the orientation vector is denoted by $\mathbf{n}_{R,j}^{(i)}$, where $j \in \{1, \dots, K_i\}$. Similarly, for the j th LED at the i th VLC unit, the location is given by $\mathbf{x}_i + \mathbf{b}_{i,j}$ and the orientation vector is represented by $\mathbf{n}_{T,j}^{(i)}$, where $j \in \{1, \dots, L_i\}$. The displacement vectors, $\mathbf{a}_{i,j}$'s and $\mathbf{b}_{i,j}$'s, are known design parameters for the VLC units. Also, the orientation vectors for the LEDs and PDs at the VLC units are assumed to be known since they can be determined by the VLC unit design and by auxiliary sensors (e.g., inertial measurement unit (IMU) consisting of gyroscope, accelerometer and magnetometer [47]–[49]⁴). To distinguish the LED transmitters at known locations from the LEDs at the VLC units, the former are called as the *LEDs on the ceiling* (as in Fig. 1) in the remainder of the text.

At a given time, each PD can communicate with a subset of all the LEDs in the system. Therefore, the following connectivity sets are defined to specify the connections between the LEDs and the PDs:

$$\tilde{S}_k^{(j)} = \{ \ell \in \{1, \dots, N_L\} \mid \ell\text{th LED on ceiling is connected to } k\text{th PD of } j\text{th VLC unit} \} \quad (1)$$

$$S_k^{(i,j)} = \{ \ell \in \{1, \dots, L_i\} \mid \ell\text{th LED of } i\text{th VLC unit is connected to } k\text{th PD of } j\text{th VLC unit} \}. \quad (2)$$

Namely, $\tilde{S}_k^{(j)}$ represents the set of LEDs on the ceiling that are connected to the k th PD at the j th VLC unit. Similarly, $S_k^{(i,j)}$ is the set of LEDs at the i th VLC unit that are connected to the k th PD at the j th VLC unit.

⁴Relative locations and orientations of PDs and LEDs on VLC units can easily be calculated via simple linear operations (i.e., rotations and translations) based on IMU orientation measurements.

The aim is to estimate the unknown locations, $\mathbf{x}_1, \dots, \mathbf{x}_{N_V}$, of the VLC units based on the RSS observations (measurements) at the PDs. Let $\tilde{P}_{\ell,k}^{(j)}$ represent the RSS observation at the k th PD of the j th VLC unit due to the transmission from the ℓ th LED on the ceiling. Similarly, let $P_{\ell,k}^{(i,j)}$ denote the RSS observation at the k th PD of the j th VLC unit due to the ℓ th LED at the i th VLC unit. Based on the Lambertian formula [12], [50], $\tilde{P}_{\ell,k}^{(j)}$ and $P_{\ell,k}^{(i,j)}$ can be expressed as follows⁵:

$$\tilde{P}_{\ell,k}^{(j)} = \tilde{\alpha}_{\ell,k}^{(j)}(\mathbf{x}_j) + \tilde{\eta}_{\ell,k}^{(j)} \quad (3)$$

$$P_{\ell,k}^{(i,j)} = \alpha_{\ell,k}^{(i,j)}(\mathbf{x}_j, \mathbf{x}_i) + \eta_{\ell,k}^{(i,j)} \quad (4)$$

where

$$\tilde{\alpha}_{\ell,k}^{(j)}(\mathbf{x}_j) \triangleq -\frac{\tilde{m}_\ell + 1}{2\pi} \tilde{P}_{T,\ell} A_k^{(j)} \frac{((\tilde{\mathbf{d}}_{\ell,k}^{(j)})^T \tilde{\mathbf{n}}_{T,\ell})^{\tilde{m}_\ell} (\tilde{\mathbf{d}}_{\ell,k}^{(j)})^T \mathbf{n}_{R,k}^{(j)}}{\|\tilde{\mathbf{d}}_{\ell,k}^{(j)}\|^{\tilde{m}_\ell + 3}} \quad (5)$$

$$\alpha_{\ell,k}^{(i,j)}(\mathbf{x}_j, \mathbf{x}_i) \triangleq -\frac{m_\ell^{(i)} + 1}{2\pi} P_{T,\ell}^{(i)} A_k^{(j)} \times \frac{((\mathbf{d}_{\ell,k}^{(i,j)})^T \mathbf{n}_{T,\ell}^{(i)})^{m_\ell^{(i)}} (\mathbf{d}_{\ell,k}^{(i,j)})^T \mathbf{n}_{R,k}^{(j)}}{\|\mathbf{d}_{\ell,k}^{(i,j)}\|^{m_\ell^{(i)} + 3}}. \quad (6)$$

for $j \in \{1, \dots, N_V\}$, $k \in \{1, \dots, K_j\}$, $i \in \{1, \dots, N_V\} \setminus j$ and $\ell \in S_k^{(i,j)}$, where $\tilde{\mathbf{d}}_{\ell,k}^{(j)} \triangleq \mathbf{x}_j + \mathbf{a}_{j,k} - \mathbf{y}_\ell$ and $\mathbf{d}_{\ell,k}^{(i,j)} \triangleq \mathbf{x}_j + \mathbf{a}_{j,k} - \mathbf{x}_i - \mathbf{b}_{i,\ell}$. In (5) and (6), \tilde{m}_ℓ ($m_\ell^{(i)}$) is the Lambertian order for the ℓ th LED on the ceiling (at the i th VLC unit), $A_k^{(j)}$ is the area of the k th PD at the j th VLC unit, and $\tilde{P}_{T,\ell}$ ($P_{T,\ell}^{(i)}$) is the transmit power of the ℓ th LED on the ceiling (at the i th VLC unit). In addition, the noise components, $\tilde{\eta}_{\ell,k}^{(j)}$ and $\eta_{\ell,k}^{(i,j)}$, are modeled by zero-mean Gaussian random variables each with a variance of $\sigma_{j,k}^2$. Considering the use of a certain multiplexing scheme (e.g., time division multiplexing among the LEDs at the same VLC unit and on the ceiling, and frequency division multiplexing among the LEDs at different VLC units or on the ceiling), $\tilde{\eta}_{\ell,k}^{(j)}$ and $\eta_{\ell,k}^{(i,j)}$ are assumed to be independent for all different (j, k) pairs and for all ℓ and i .

B. ML Estimator and CRLB

Let $\mathbf{x} \triangleq [\mathbf{x}_1^T \dots \mathbf{x}_{N_V}^T]^T$ denote the vector of unknown parameters (which has a size of $3N_V \times 1$) and let \mathbf{P} represent a vector consisting of all the measurements in (3) and (4). The elements of \mathbf{P} can be expressed as follows:

$$\left\{ \left\{ \left\{ \tilde{P}_{\ell,k}^{(j)} \right\}_{\ell \in \tilde{S}_k^{(j)}} \right\}_{k \in \{1, \dots, K_j\}} \right\}_{j \in \{1, \dots, N_V\}},$$

$$\left\{ \left\{ \left\{ \left\{ P_{\ell,k}^{(i,j)} \right\}_{\ell \in S_k^{(i,j)}} \right\}_{i \in \{1, \dots, N_V\} \setminus \{j\}} \right\}_{k \in \{1, \dots, K_j\}} \right\}_{j \in \{1, \dots, N_V\}}.$$

Then, the conditional probability density function (PDF) of \mathbf{P} given \mathbf{x} , i.e., the likelihood function, can be stated as

$$f(\mathbf{P} | \mathbf{x}) = \left(\prod_{j=1}^{N_V} \prod_{k=1}^{K_j} \frac{1}{(\sqrt{2\pi} \sigma_{j,k})^{N_{tot}^{(j,k)}}} \right) e^{-\sum_{j=1}^{N_V} \sum_{k=1}^{K_j} \frac{h_{j,k}(\mathbf{x})}{2\sigma_{j,k}^2}} \quad (7)$$

where $N_{tot}^{(j,k)}$ represents the total number of LEDs that can communicate with the k th PD at the j th VLC unit; that is, $N_{tot}^{(j,k)} \triangleq |\tilde{S}_k^{(j)}| + \sum_{i=1, i \neq j}^{N_V} |S_k^{(i,j)}|$, and $h_{j,k}(\mathbf{x})$ is defined as

⁵Since the wavelength of the infrared/visible light carrier (on the order of 10^{-6} m) is much lower than dimensions of typical PDs (i.e., 10^{-2} m), multipath reflections are averaged out by integration of the incident optical power over the area of a PD [33], [36], [51]. Hence, in this study, only line-of-sight (LOS) links are taken into account in the infrared/visible light channel model.

$$h_{j,k}(\mathbf{x}) \triangleq \sum_{\ell \in \tilde{S}_k^{(j)}} (\tilde{P}_{\ell,k}^{(j)} - \tilde{\alpha}_{\ell,k}^{(j)}(\mathbf{x}_j))^2 + \sum_{i=1, i \neq j}^{N_V} \sum_{\ell \in S_k^{(i,j)}} (P_{\ell,k}^{(i,j)} - \alpha_{\ell,k}^{(i,j)}(\mathbf{x}_j, \mathbf{x}_i))^2. \quad (8)$$

From (7), the maximum likelihood estimator (MLE) is obtained as

$$\hat{\mathbf{x}}_{\text{ML}} = \arg \min_{\mathbf{x}} \sum_{j=1}^{N_V} \sum_{k=1}^{K_j} \frac{h_{j,k}(\mathbf{x})}{\sigma_{j,k}^2} \quad (9)$$

and the Fisher information matrix (FIM) [52] is given by

$$[\mathbf{J}]_{t_1, t_2} = \mathbb{E} \left\{ \frac{\partial \log f(\mathbf{P} | \mathbf{x})}{\partial x_{t_1}} \frac{\partial \log f(\mathbf{P} | \mathbf{x})}{\partial x_{t_2}} \right\} \quad (10)$$

where x_{t_1} (x_{t_2}) represents element t_1 (t_2) of vector \mathbf{x} with $t_1, t_2 \in \{1, 2, \dots, 3N_V\}$. Then, the CRLB is stated as

$$\text{CRLB} = \text{trace}(\mathbf{J}^{-1}) \leq \mathbb{E}\{\|\hat{\mathbf{x}} - \mathbf{x}\|^2\} \quad (11)$$

where $\hat{\mathbf{x}}$ represents an unbiased estimator of \mathbf{x} . From (7) and (8), the elements of the FIM in (10) can be calculated after some manipulation as

$$[\mathbf{J}]_{t_1, t_2} = \sum_{j=1}^{N_V} \sum_{k=1}^{K_j} \frac{1}{\sigma_{j,k}^2} \left(\sum_{\ell \in \tilde{S}_k^{(j)}} \frac{\partial \tilde{\alpha}_{\ell,k}^{(j)}(\mathbf{x}_j)}{\partial x_{t_1}} \frac{\partial \tilde{\alpha}_{\ell,k}^{(j)}(\mathbf{x}_j)}{\partial x_{t_2}} + \sum_{i=1, i \neq j}^{N_V} \sum_{\ell \in S_k^{(i,j)}} \frac{\partial \alpha_{\ell,k}^{(i,j)}(\mathbf{x}_j, \mathbf{x}_i)}{\partial x_{t_1}} \frac{\partial \alpha_{\ell,k}^{(i,j)}(\mathbf{x}_j, \mathbf{x}_i)}{\partial x_{t_2}} \right). \quad (12)$$

Based on (11) and (12), the CRLB for location estimation can be obtained for cooperative VLP networks (please see Section S-VIII in the supplementary material for the partial derivatives in (12)). The obtained CRLB expression is generic for any three-dimensional configuration and covers all possible cooperation scenarios via the definitions of the connectivity sets (see (1) and (2)). To the best of authors' knowledge, such a CRLB expression has not been available in the literature for cooperative VLP networks.

Remark 1: From (12), it is noted that the first summation term in the parentheses is related to the information from the LED transmitters on the ceiling whereas the remaining terms are due to the cooperation among the VLC units. In the noncooperative case, the elements of the FIM are given by the expression in the first line of (12).

Via (11) and (12), the effects of cooperation on the accuracy of VLP networks can be quantified, as investigated in Section VI.

Remark 2: According to [53], the irradiance value of the infrared emitters should be lower than 100 W/m^2 to eliminate the cornea/lens risk in the range from 780 nm to 3000 nm . Based on [36, Sec. II-C], the irradiance at a distance d (m) and an irradiation angle ϕ is given by

$$\mathcal{I}(d, \phi) = P_T \frac{m+1}{2\pi d^2} \cos^m(\phi) \quad (\text{W/m}^2) \quad (13)$$

where P_T is the optical power of the LED (W) and m is the Lambertian order. Setting $m = 1$ and $\phi = 0$ (i.e., directly looking at the infrared source, which is the worst case from an eye-safety perspective), the value in (13) becomes

$$\mathcal{I}(d, \phi = 0^\circ) = \frac{P_T}{\pi d^2} \quad (\text{W/m}^2). \quad (14)$$

Considering the limit value of 100 W/m^2 , for an eye looking directly at the infrared source at a distance of $d = 0.2 \text{ m}$, the optical power of the infrared LED should satisfy

$$P_T \leq 12.56 \text{ W} \quad (15)$$

so that the LED does not pose any risks to the cornea/lens of the eye. As observed from the numerical results in Section VI, significant cooperation benefits can be achieved with 1 W optical power for the LEDs on the VLC units (i.e., the LEDs employed for cooperation). Therefore, the optical power values for which the proposed system can enjoy the advantages of cooperation are well below the eye-safety limits of the cornea/lens presented in [53]. Regarding the thermal retinal risks and blue light hazard, the irradiance limit is given by $3.77 \text{ W/sr}/(0.2)^2 = 94.25 \text{ W/m}^2$ [53], which yields the condition $P_T \leq 11.84 \text{ W}$ for $d = 0.2 \text{ m}$. Similar to the case of the cornea/lens risk, the optical power values in typical indoor scenarios considered in this paper satisfy the eye-safety limits related to the thermal retinal risks and blue light hazard.

III. COOPERATIVE LOCALIZATION AS A QUASICONVEX FEASIBILITY PROBLEM

In this section, the problem of cooperative localization in VLP networks is investigated in the framework of convex/quasiconvex feasibility. First, the feasibility approach to the localization problem is motivated, and the problem formulation is presented. Then, the convexity analysis is carried out for the resulting constraint sets.

A. Motivation

For the localization of the VLC units, the MLE in (9) has very high computational complexity as it requires a search over a $3N_V$ dimensional space. In addition, the formulation in (9) presents a nonconvex optimization problem; hence, convex optimization tools cannot be employed to obtain the (global) optimal solution of (9). As the number of VLC units increases, centralized approaches obtained as solutions to a given optimization problem (such as (9)) may become computationally prohibitive. Besides scalability issues, centralized methods also require all measurements gathered at the VLC units to be relayed to a central unit for joint processing, which may lead to communication bottlenecks. Therefore, low-complexity algorithms amenable to distributed implementation are needed to efficiently solve the cooperative localization problem in VLP networks. To that aim, the localization problem is cast as a feasibility problem with the purpose of finding a point in a finite dimensional Euclidean space that lies within the intersection of some constraint sets. Feasibility-seeking methods enjoy the advantage of not requiring an objective function, thereby eliminating the concerns for nonconvexity or nondifferentiability of the objective function [54]. Hence, modeling the localization problem as a feasibility problem (i) alleviates the computational burden of minimizing a (possibly nonconvex) cost function in the highly unfavorable centralized setting and (ii) facilitates the use of efficient distributed algorithms involving parallel or sequential processing at individual VLC units.

B. Problem Formulation

Considering the Lambertian formula in (3)–(6), an RSS measurement at a PD can be expressed as

$$\hat{P}_r = P_r + \eta \quad (16)$$

where P_r is the true observation (as in (5) or (6)) and η is the measurement noise. Suppose that the RSS measurement errors

are negative, which yields $\hat{P}_r \leq P_r$.⁶ Then, based on (5) and (6), the following inequality is obtained:

$$g(\mathbf{x}; \mathbf{y}, \mathbf{n}_T, \mathbf{n}_R, m, \gamma) \leq 0 \quad (17)$$

where $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is the Lambertian function with respect to the unknown PD location \mathbf{x} , defined as

$$g(\mathbf{x}; \mathbf{y}, \mathbf{n}_T, \mathbf{n}_R, m, \gamma) \triangleq \gamma - \frac{[(\mathbf{x} - \mathbf{y})^T \mathbf{n}_T]^m (\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^{m+3}}, \quad (18)$$

\mathbf{y} , \mathbf{n}_T , \mathbf{n}_R , and m are known, d is the dimension of the visible light localization network, and γ is given by $\gamma = \frac{\hat{P}_r}{P_t} \frac{2\pi}{(m+1)A}$. The field-of-views (FOVs) of the LED transmitters and the PDs are taken as 90° , which implies that $(\mathbf{x} - \mathbf{y})^T \mathbf{n}_T \geq 0$ and $(\mathbf{y} - \mathbf{x})^T \mathbf{n}_R \geq 0$. Under the assumption of negative measurement errors, the feasible set in which the true PD location resides is given by the following lower level set of $g(\mathbf{x})$:

$$\mathcal{L} = \left\{ \mathbf{x} \in \mathbb{R}^d \mid g(\mathbf{x}; \mathbf{y}, \mathbf{n}_T, \mathbf{n}_R, m, \gamma) \leq 0 \right\} \quad (19)$$

which will hereafter be referred to as the *Lambertian set*. In RF wireless localization networks, such feasible sets are generally obtained as balls [27], [29], hyperplanes [56], or ellipsoids [57], all of which lead to closed-form expressions for orthogonal projection. For $k \in \{1, 2, \dots, K_j\}$ and $j \in \{1, 2, \dots, N_V\}$, the Lambertian set corresponding to the k th PD of the j th VLC unit based on the signal received from the ℓ th LED on the ceiling for $\ell \in \tilde{S}_k^{(j)}$ is defined as follows:

$$\mathcal{N}_{\ell,k}^{(j)} = \left\{ \mathbf{z} \in \mathbb{R}^d \mid \tilde{g}_{\ell,k}^{(j)}(\mathbf{z}) \leq 0 \right\} \quad (20)$$

where $\tilde{g}_{\ell,k}^{(j)}(\mathbf{z})$ is given by

$$\tilde{g}_{\ell,k}^{(j)}(\mathbf{z}) \triangleq g(\mathbf{z}; \mathbf{y}_\ell - \mathbf{a}_{j,k}, \tilde{\mathbf{n}}_{T,\ell}, \mathbf{n}_{R,k}^{(j)}, \tilde{m}_\ell, \tilde{\gamma}_{\ell,k}^{(j)}) \quad (21)$$

and $\tilde{\gamma}_{\ell,k}^{(j)}$ is calculated from (3). Similarly, the Lambertian set corresponding to the k th PD of the j th VLC unit based on the signal received from the ℓ th LED of the i th VLC unit for $\ell \in S_k^{(i,j)}$ is defined as

$$\mathcal{C}_{\ell,k}^{(i,j)} = \left\{ \mathbf{z} \in \mathbb{R}^d \mid g_{\ell,k}^{(i,j)}(\mathbf{z}, \mathbf{x}_i) \leq 0 \right\} \quad (22)$$

where $g_{\ell,k}^{(i,j)}(\mathbf{z}, \mathbf{x}_i)$ is given by

$$g_{\ell,k}^{(i,j)}(\mathbf{z}, \mathbf{x}_i) \triangleq g(\mathbf{z}; \mathbf{x}_i + \mathbf{b}_{i,\ell} - \mathbf{a}_{j,k}, \mathbf{n}_{T,\ell}^{(i)}, \mathbf{n}_{R,k}^{(j)}, m_\ell^{(i)}, \gamma_{\ell,k}^{(i,j)}) \quad (23)$$

and $\gamma_{\ell,k}^{(i,j)}$ is calculated from (4). The sets defined as in (20) represent noncooperative localization as they are constructed from the RSS measurements corresponding to the LEDs on the ceiling, whereas the sets in (22) are based on the signals from the LEDs of the other VLC units and represent the cooperation among the VLC units. Assuming negatively biased RSS measurements, the problem of cooperative localization in a visible light network reduces to that of finding a point in the intersection of sets as defined in (20) and (22) for each VLC unit. If the Lambertian function in (18) is assumed to be

⁶In order to satisfy the negative error assumption, a constant value can always be subtracted from the actual RSS measurement [55]. Decreasing the value of an RSS measurement is equivalent to enlarging the corresponding feasible set. Although this assumption does not have a physical justification, it facilitates theoretical derivations and feasibility modeling of the localization problem. It will be justified via simulations in Section VI-B that the proposed feasibility-seeking algorithms will converge for realistic noise models (e.g., Gaussian), as well.

quasiconvex⁷, then the *quasiconvex feasibility problem* (QFP) can be formulated as follows [37], [58]:

Problem 1: Let $\mathbf{x} \triangleq (\mathbf{x}_1, \dots, \mathbf{x}_{N_V})$. The feasibility problem for cooperative localization of VLC units is given by⁸

$$\begin{aligned} & \text{find } \mathbf{x} \in \mathbb{R}^{dN_V} \\ & \text{subject to } \mathbf{x}_j \in \Lambda_j \cap \Upsilon_j, j = 1, \dots, N_V \end{aligned} \quad (24)$$

where

$$\Lambda_j = \bigcap_{k=1}^{K_j} \bigcap_{\ell \in \tilde{S}_k^{(j)}} \mathcal{N}_{\ell,k}^{(j)} \quad (25)$$

$$\Upsilon_j = \bigcap_{k=1}^{K_j} \bigcap_{i=1}^{N_V} \bigcap_{\ell \in S_k^{(i,j)}} \mathcal{C}_{\ell,k}^{(i,j)}. \quad (26)$$

C. Convexity Analysis of Lambertian Sets

The Lambertian sets as defined in (19) are not convex in general. The following lemma presents the conditions under which the Lambertian sets become convex.

Lemma 1: Consider the α -sublevel set

$$\mathcal{B}_\alpha = \left\{ \mathbf{x} \in \Omega \mid g_\epsilon(\mathbf{x}) \leq \alpha \right\} \quad (27)$$

of $g_\epsilon(\mathbf{x})$, which is given by

$$g_\epsilon(\mathbf{x}) = \gamma - \frac{(\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^k + \epsilon} \quad (28)$$

where ϵ is a small positive constant to avoid non-differentiability and non-continuity of $g_\epsilon(\cdot)$ at \mathbf{y} , as in [59, Eq. 7], $k \geq 1$ and $\gamma > 0$ are real numbers, and $\Omega \subset \mathbb{R}^d$ is defined as

$$\Omega = \left\{ \mathbf{x} \in \mathbb{R}^d \mid (\mathbf{y} - \mathbf{x})^T \mathbf{n}_R \geq 0 \right\}. \quad (29)$$

Then, \mathcal{B}_α is convex for each $\alpha \in \mathbb{R}$.

Proof: Please see Section S-I in the supplementary material.

Remark 3: Lemma 1 characterizes the type of Lambertian functions whose sublevel sets are convex. Since a function whose all sublevel sets are convex is quasiconvex [60], Lambertian functions of the form (28) are quasiconvex over the halfspace Ω in (29). It can be noted that Ω consists of those VLC unit locations which are able to obtain measurements from an LED located at \mathbf{y} due to the receiver FOV limit of 90° .

D. Convexification of Lambertian Sets

In this part, we utilize Lemma 1 to investigate the following two cases in which the Lambertian functions can be transformed into the form of (28) and Problem 1 becomes a QFP.

⁷The conditions under which the Lambertian function is quasiconvex are investigated in Section III-C.

⁸It may be more convenient to regard the problem in (24) as an *implicit* quasiconvex feasibility problem (IQFP) since the Lambertian sets $\mathcal{C}_{\ell,k}^{(i,j)}$ depend on the locations of the VLC units, which are not known *a priori* [29]. It should be emphasized that the feasibility problem posed in Problem 1 is different from those in RF-based localization systems (e.g., [28], [29]) since the constraint sets and the associated quasiconvex functions have distinct characteristics as compared to convex functions (e.g., distance to a ball) encountered in RF-based systems.

1) *Case 1: Convexification via Majorization:* We propose to approximate the Lambertian function $g(\mathbf{x})$ in (18) by a *quasiconvex minorant* $\tilde{g}(\mathbf{x})$ such that $\tilde{g}(\mathbf{x}) \leq g(\mathbf{x})$ for $\mathbf{x} \in \Omega$ and $\mathcal{L} \subseteq \tilde{\mathcal{L}}$, where $\tilde{\mathcal{L}} \triangleq \{\mathbf{x} \in \Omega \mid \tilde{g}(\mathbf{x}) \leq 0\}$ represents a majorization of the original set $\mathcal{L} \triangleq \{\mathbf{x} \in \Omega \mid g(\mathbf{x}) \leq 0\}$. Assuming $\mathbf{x} \in \Omega$, we have

$$g(\mathbf{x}) = \gamma - \frac{[(\mathbf{x} - \mathbf{y})^T \mathbf{n}_T]^m (\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^{m+3}} \quad (30)$$

$$\geq \gamma - \frac{\|\mathbf{x} - \mathbf{y}\|^m \|\mathbf{n}_T\|^m (\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^{m+3}} \quad (31)$$

$$= \gamma - \frac{(\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^3} \triangleq \tilde{g}(\mathbf{x}) \quad (32)$$

where (31) is due to the Cauchy-Schwarz inequality and $\mathbf{x} \in \Omega$, and (32) follows from the unit norm property of the orientation vector. Then, including ϵ in the denominator, we construct the Lambertian sets as (hereafter called *expanded Lambertian sets*)

$$\mathcal{L} = \left\{ \mathbf{x} \in \Omega \mid \tilde{g}_\epsilon(\mathbf{x}) \leq 0 \right\} \quad (33)$$

with

$$\tilde{g}_\epsilon(\mathbf{x}) = \gamma - \frac{(\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^3 + \epsilon} \quad (34)$$

and Ω being as in (29). According to Lemma 1, \mathcal{L} in (33) is convex, $\tilde{g}_\epsilon(\mathbf{x})$ in (34) is quasiconvex over Ω and the resulting problem of determining a point inside the intersection of such sets turns into a QFP, which can be studied through iterative projection algorithms [37], [58].

2) *Case 2: Known VLC Height, Perpendicular LED:* In this case, as in [10], [12], [13], [61], it is assumed that the LED transmitters on the ceiling have perpendicular orientations, i.e., $\tilde{\mathbf{n}}_{T,j} = [0 \ 0 \ -1]^T$ for each $j \in \{1, 2, \dots, N_L\}$, and the height of each VLC unit is known. This assumption is valid for some practical scenarios, an example of which is a VLP network where the LEDs on the ceiling are pointing downwards and the VLC units are attached to robots that move over a two-dimensional plane [7, Fig. 3]. Assuming that the height of the LED transmitters relative to the VLC units is h and $\mathbf{n}_T = [0 \ 0 \ -1]^T$, the Lambertian function in (18) can be rewritten as follows:

$$g(\mathbf{x}; \mathbf{y}, \mathbf{n}_T, \mathbf{n}_R, m, \gamma) = \gamma - \frac{h^m (\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^{m+3}}. \quad (35)$$

Then, the Lambertian set corresponding to the function in (35) by introducing ϵ in the denominator is obtained as

$$\mathcal{L} = \left\{ \mathbf{x} \in \Omega \mid \tilde{g}_\epsilon(\mathbf{x}) \leq 0 \right\} \quad (36)$$

with

$$\tilde{g}_\epsilon(\mathbf{x}) = \tilde{\gamma} - \frac{(\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^{m+3} + \epsilon} \quad (37)$$

where Ω is given by (29) and $\tilde{\gamma} = \gamma/h^m$. Note that the Lambertian set in (36) is effectively defined on \mathbb{R}^2 since the height of the VLC unit is already known. According to Lemma 1, the set defined in (36) is convex. Therefore, in this case, the noncooperative sets as defined in (20) are originally convex.

Based on the discussion above, it is concluded that in the case of a known VLC height and perpendicular LED transmitter orientations, the expanded Lambertian sets in (33) defined on \mathbb{R}^2 must be used for the measurements among the VLC units (i.e., cooperative measurements) in order to ensure that Problem 1

is a QFP. For the general case in which the LED orientations are arbitrary and/or the heights of the VLC units are unknown, all the noncooperative and cooperative Lambertian sets must be replaced by the corresponding expanded versions in (33).

A noncooperative VLP network is illustrated in Fig. 2(a), where there exist four LED transmitters on the ceiling and two VLC units. In the network, it is assumed that the heights of the VLC units are known and the LEDs on the ceiling have perpendicular orientations so that Case 2 type convex Lambertian sets can be utilized for the measurements between the LEDs on the ceiling and the VLC units. Fig. 2(b) shows the cooperative version of the VLP network with cooperative Lambertian sets including both the nonexpanded (original) sets as in (22) and Case 1 type expanded sets as in (33). It is noted from Fig. 2(b) that incorporating cooperative Lambertian sets into the localization geometry can significantly reduce the region of intersection of the Lambertian sets.

IV. GRADIENT PROJECTIONS ALGORITHMS

In this section, we design iterative subgradient projections based algorithms to solve Problem 1. The idea of using subgradient projections is to approach a convex set defined as a lower contour set of a convex/quasiconvex function by moving in the direction that decreases the value of that function at each iteration, i.e., in the opposite direction of the subgradient of the function at the current iterate [44], [62]. First, the definition of the gradient projector is presented as follows:

Definition 1: The gradient projection operator $G_f^\lambda : \mathbb{R}^d \rightarrow \mathbb{R}^d$ onto the zero-sublevel set of a continuously differentiable function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is given by [63]

$$G_f^\lambda : \mathbf{x} \mapsto \begin{cases} \mathbf{x} - \lambda \frac{f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|^2} \nabla f(\mathbf{x}), & \text{if } f(\mathbf{x}) > 0 \\ \mathbf{x}, & \text{if } f(\mathbf{x}) \leq 0 \end{cases} \quad (38)$$

where λ is the relaxation parameter and ∇ is the gradient operator. The gradient projector can also be expressed as

$$G_f^\lambda(\mathbf{x}) = \mathbf{x} - \lambda \frac{f^+(\mathbf{x})}{\|\nabla f(\mathbf{x})\|^2} \nabla f(\mathbf{x}) \quad (39)$$

with $f^+(\mathbf{x})$ denoting the positive part, i.e., $f^+(\mathbf{x}) = \max\{0, f(\mathbf{x})\}$. In the sequel, it is assumed that $G_f^\lambda(\mathbf{x}) = \mathbf{x}$ when \mathbf{x} is outside the region where f is quasiconvex.

A. Projection Onto Intersection of Halfspaces

Since the functions of the form (28) are continuously differentiable and quasiconvex on the halfspace Ω in (29), a special case of subgradient projections, namely, gradient projections, can be utilized to solve Problem 1, under the constraint that iterates must be inside Ω to guarantee quasiconvexity. Hence, at the start of each iteration of gradient projections, projections onto the intersection of halfspaces of the form Ω in (29) can be performed to keep the iterates inside the quasiconvex region. The procedure for projection onto the intersection of halfspaces

$$\Gamma_j = \bigcap_{k=1}^{K_j} \bigcap_{\ell \in \tilde{\mathcal{S}}_k^{(j)}} \tilde{\Omega}_{\ell,k}^{(j)} \quad (40)$$

corresponding to the j th VLC unit for $j \in \{1, 2, \dots, N_V\}$, with the halfspaces given by

$$\tilde{\Omega}_{\ell,k}^{(j)} \triangleq \left\{ \mathbf{x} \in \mathbb{R}^d \mid (\mathbf{y}_\ell - \mathbf{a}_{j,k} - \mathbf{x})^T \mathbf{n}_{R,k}^{(j)} \geq 0 \right\}, \quad (41)$$

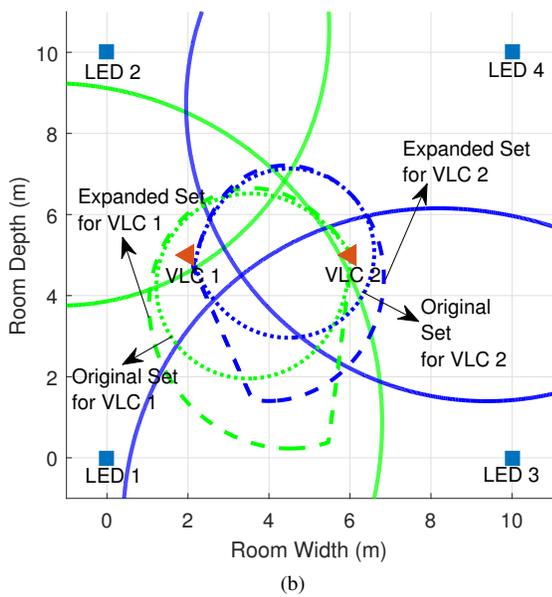
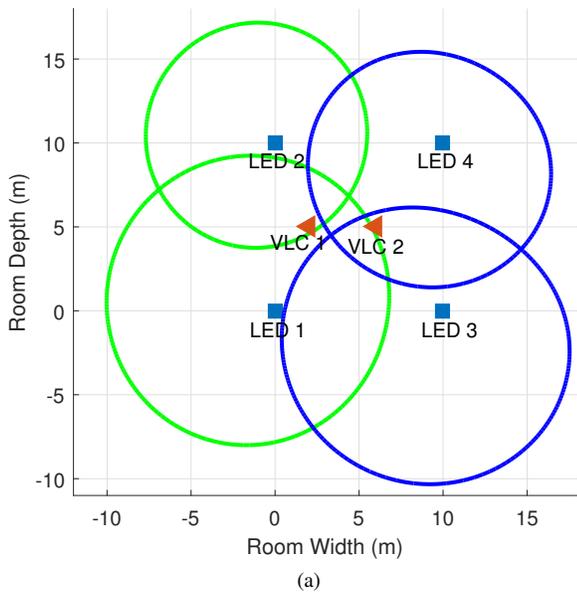


Fig. 2. (a) A noncooperative VLP network consisting of four LED transmitters on ceiling and two VLC units. VLC-1 is connected to LED-1 and LED-2, and VLC-2 is connected to LED-3 and LED-4. Green and blue regions represent the noncooperative Lambertian sets for VLC-1 and VLC-2, respectively. (b) Cooperative version of the VLP network in Fig. 2(a), shown by zooming onto VLC units. Case 1 type expanded cooperative Lambertian sets and their nonexpanded (original) counterparts are illustrated along with noncooperative Lambertian sets. Cooperation helps shrink the intersection region of Lambertian sets for VLC units.

is provided in Algorithm 1⁹. In order to find a point inside the intersection of halfspaces, the method of alternating (cyclic) projections is employed in Algorithm 1, where the current iterate is projected onto each halfspace in a cyclic manner. Convergence properties of this method are well studied in the literature [64], [65]. Γ_j is guaranteed to be nonempty since it represents the set of possible locations for the j th VLC unit at which the RSS measurements from the connected LEDs on the ceiling can be acquired. However, the intersection of the halfspaces corresponding to the LEDs of the other VLC units that are

⁹ $P_C(\mathbf{x})$ denotes the orthogonal projection operator, i.e., $P_C(\mathbf{x}) = \arg \min_{\mathbf{w} \in C} \|\mathbf{w} - \mathbf{x}\|$.

connected to the j th VLC unit may be empty due to the VLC unit locations being unknown and variable during iterations.

Algorithm 1 Projection Onto Intersection of Halfspaces Γ_j

```

function  $P_{\Gamma_j}(\mathbf{x}_j)$ 
  Initialization:  $\mathbf{x}_j^{(0)} = \mathbf{x}_j$ 
  Iterative Step: Given the  $n$ th iterate  $\mathbf{x}_j^{(n)} \in \mathbb{R}^d$ 
  for  $k = 1, \dots, K_j$  do
    for  $\ell \in \tilde{S}_k^{(j)}$  do
       $\mathbf{x}_j^{(n)} = P_{\tilde{\Omega}_{\ell,k}^{(j)}}(\mathbf{x}_j^{(n)})$ 
    end for
  end for
  Set  $\mathbf{x}_j^{(n+1)} = \mathbf{x}_j^{(n)}$ 
  Stopping Criterion:  $\|\mathbf{x}_j^{(n+1)} - \mathbf{x}_j^{(n)}\| < \delta$  for some  $\delta > 0$ .
end function

```

B. Step Size Selection

An important phase of the proposed projection algorithms is determining the relaxation parameters (i.e., step sizes) associated with the gradient projector. The step size selection procedure exploits the well-known Armijo rule, which is an inexact line search method used extensively for gradient descent methods in the literature [66], [67], [68, Section 1.2]. Algorithm 2 provides an Armijo-like procedure for step size selection given a set of Lambertian functions, the initial step size value λ , a fixed constant $\beta \in (0, 1)$ specifying the degree of decline in the value of the function, step size shrinkage factor $\xi \in (0, 1)$, and the current point \mathbf{x} . The guarantee of existence of a step size as described in Algorithm 2 can be shown similarly to [69, Lemma 4].

Algorithm 2 Armijo Rule for Step Size Selection

```

function  $\mathcal{J}(\{f_i\}_{i=1}^M, \lambda, \beta, \xi, \mathbf{x})$ 
  Output: New step size  $\tilde{\lambda}$ 
  Set the step size as
     $\tilde{\lambda} = \lambda \xi^{\tilde{m}}$ 
  where
     $\tilde{m} = \min\{m \in \mathbb{Z}_{\geq 0} \mid$ 
       $f_i(G_{f_i}^{\lambda \xi^m}(\mathbf{x})) \leq f_i(\mathbf{x})(1 - \beta \lambda \xi^m), \forall i \in \{1, 2, \dots, M\}\}$ 
    end function

```

C. Iterative Projection Based Algorithms

In this work, two classes of gradient projections algorithms, namely, sequential (i.e., cyclic) [62] and simultaneous (i.e., parallel) [70] projections, are considered for the QFP described in Problem 1. The proposed algorithm for cyclic projections, namely, the cooperative cyclic gradient projections (CCGP) algorithm, for cooperative localization of VLC units is provided in Algorithm 3. In the proposed cyclic projections, the current iterate, which signifies the location of the given VLC unit, is first projected onto the intersection of halfspaces corresponding to the LEDs on the ceiling via Algorithm 1. Then, the resulting point is projected onto the noncooperative Lambertian set that leads to the highest function value, i.e., the most violated constraint set [37]. Similarly, projection onto the most violated constraint set among the cooperative Lambertian sets is performed and the projections obtained by noncooperative and cooperative sets are weighted to obtain the next iterate.

The cooperative simultaneous gradient projections (CSGP) algorithm is proposed as detailed in Algorithm 4. Simultaneous projections are based on projecting the current point onto each noncooperative and cooperative Lambertian set separately and then averaging all the resulting points to obtain the next iterate.

At each iteration, the parallel projection stage is preceded by projection onto the intersection of halfspaces, which aims to ensure that the current iterate resides in the region where all the Lambertian functions corresponding to the fixed anchors (i.e., the LEDs on the ceiling) are quasiconvex. It should be noted that for both cyclic and simultaneous projections, the cooperative Lambertian sets are determined by the latest estimates of the VLC unit locations [29], which are updated in the ascending order of their indices. In addition, the step sizes are updated using the Armijo rule in Algorithm 2.

Remark 4: Both Algorithm 3 and Algorithm 4 can be implemented in a distributed manner by employing a gossip-like procedure among the VLC units [71]. After refining its location estimate via projection methods, each VLC unit broadcasts the resulting updated location to other VLC units to which it is connected. In order to save computation time, a synchronous counterpart of this asynchronous/sequential algorithm can be devised, where VLC units work in parallel to update their locations based on the most recent broadcast information. Hence, the synchronous/parallel implementation trades off the localization accuracy for faster convergence to the desired solution.

Algorithm 3 Cooperative Cyclic Gradient Projections (CCGP)

Initialization: Choose an arbitrary initial point $(\mathbf{x}_1^{(0)}, \dots, \mathbf{x}_{N_V}^{(0)}) \in \mathbb{R}^{dN_V}$.

Iterative Step: Given the n th iterate $(\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{N_V}^{(n)}) \in \mathbb{R}^{dN_V}$

for $j = 1, \dots, N_V$ **do**

Projection Onto Intersection of Halfspaces Γ_j by Algorithm 1:

$$\tilde{\mathbf{x}}_j^{(n)} = P_{\Gamma_j}(\mathbf{x}_j^{(n)}) \quad (45)$$

Most Violated Constraint Control for Noncooperative Projections:

$$(\hat{k}_{nc}, \hat{\ell}_{nc}) = \arg \max_{k, \ell} \tilde{g}_{\ell, k}^{(j)}(\tilde{\mathbf{x}}_j^{(n)}) \quad (46)$$

Most Violated Constraint Control for Cooperative Projections:

$$(\hat{k}_c, \hat{i}_c, \hat{\ell}_c) = \arg \max_{k, i, \ell} \mathcal{G}_j^{(n)} \quad (47)$$

where $\mathcal{G}_j^{(n)} \triangleq \left\{ g_{\ell, k}^{(i, j)}(\tilde{\mathbf{x}}_j^{(n)}, \mathbf{x}_i^{(\hat{n})}) \mid \tilde{\mathbf{x}}_j^{(n)} \in \Omega_{\ell, k}^{(i, j)} \right\}$ (48)

$$\Omega_{\ell, k}^{(i, j)} \triangleq \left\{ \mathbf{x} \in \mathbb{R}^d \mid (\mathbf{x}_i^{(\hat{n})} + \mathbf{b}_{i, \ell} - \mathbf{a}_{j, k} - \mathbf{x})^T \mathbf{n}_{R, k}^{(j)} \geq 0 \right\} \quad (49)$$

with $\hat{n} = n$ for $i > j$, $\hat{n} = n + 1$ for $i < j$.

Averaging:

$$\mathbf{x}_j^{(n+1)} = \vartheta_{nc} G_{\tilde{g}_{\ell_{nc}, \hat{k}_{nc}}^{(j)}}^{\lambda_{j, nc}^{(n)}}(\tilde{\mathbf{x}}_j^{(n)}) + \vartheta_c G_{g_{\ell_c, \hat{k}_c}^{(\hat{i}_c, j)}(\cdot, \mathbf{x}_i^{(\hat{n})})}^{\lambda_{j, c}^{(n)}}(\tilde{\mathbf{x}}_j^{(n)}) \quad (50)$$

where $\vartheta_{nc} + \vartheta_c = 1$ and $\vartheta_{nc} \geq 0$, $\vartheta_c \geq 0$.

end for

Stopping Criterion: $\sum_{j=1}^{N_V} \|\mathbf{x}_j^{(n+1)} - \mathbf{x}_j^{(n)}\|^2 < \delta$ for some $\delta > 0$.

Relaxation Parameters: Initialize $\lambda_{j, nc}^{(0)} = \lambda_{j, c}^{(0)} = \lambda_0$ and update using Algorithm 2 as

$$\lambda_{j, nc}^{(n)} = \mathcal{J}(\tilde{g}_{\ell_{nc}, \hat{k}_{nc}}^{(j)}, \lambda_{j, nc}^{(n-1)}, \beta, \xi, \tilde{\mathbf{x}}_j^{(n)}) \quad (51)$$

$$\lambda_{j, c}^{(n)} = \begin{cases} \mathcal{J}(g_{\ell_c, \hat{k}_c}^{(\hat{i}_c, j)}(\cdot, \mathbf{x}_i^{(\hat{n})}), \lambda_{j, c}^{(n-1)}, \beta, \xi, \tilde{\mathbf{x}}_j^{(n)}), & \text{if } \mathcal{G}_j^{(n)} \neq \emptyset \\ \lambda_{j, c}^{(n-1)} & \text{otherwise} \end{cases} \quad (52)$$

for $j \in \{1, 2, \dots, N_V\}$.

D. Complexity Analysis

In this part, we provide the complexity analysis of Algorithm 3, Algorithm 4, and the MLE in (9).

1) *Complexity Analysis for Algorithm 3:* We first analyze the complexity of Algorithm 3 for the j th VLC unit at each iteration. Each sub-step in the iterative step of Algorithm 3 is investigated as follows:

Assume that Algorithm 1 requires $\mathcal{O}(N_1)$ iterations for convergence. Then, the computational complexity of (45) is

Algorithm 4 Cooperative Simultaneous Gradient Projections (CSGP)

Initialization: Choose an arbitrary initial point $(\mathbf{x}_1^{(0)}, \dots, \mathbf{x}_{N_V}^{(0)}) \in \mathbb{R}^{dN_V}$.

Iterative Step: Given the n th iterate $(\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{N_V}^{(n)}) \in \mathbb{R}^{dN_V}$

for $j = 1, \dots, N_V$ **do**

Projection Onto Intersection of Halfspaces Γ_j by Algorithm 1:

$$\tilde{\mathbf{x}}_j^{(n)} = P_{\Gamma_j}(\mathbf{x}_j^{(n)}) \quad (53)$$

Parallel Projection Onto Lambertian Sets:

$$\mathbf{x}_j^{(n+1)} = \sum_{k=1}^{K_j} \left[\sum_{\ell \in \tilde{S}_k^{(j)}} \tilde{\kappa}_{\ell, k}^{(j)} G_{\tilde{g}_{\ell, k}^{(j)}}^{\lambda_{j, k}^{(n)}}(\tilde{\mathbf{x}}_j^{(n)}) + \sum_{i=1, i \neq j}^{N_V} \sum_{\ell \in S_k^{(i, j)}} \kappa_{\ell, k}^{(i, j)} G_{g_{\ell, k}^{(i, j)}(\cdot, \mathbf{x}_i^{(\hat{n})})}^{\lambda_{j, k}^{(n)}}(\tilde{\mathbf{x}}_j^{(n)}) \right] \quad (54)$$

where $\hat{n} = n$ for $i > j$, $\hat{n} = n + 1$ for $i < j$ and the weights satisfy

$$\sum_{k=1}^{K_j} \left(\sum_{\ell \in \tilde{S}_k^{(j)}} \tilde{\kappa}_{\ell, k}^{(j)} + \sum_{i=1, i \neq j}^{N_V} \sum_{\ell \in S_k^{(i, j)}} \kappa_{\ell, k}^{(i, j)} \right) = 1 \quad (55)$$

and $\tilde{\kappa}_{\ell, k}^{(j)} \geq 0$, $\kappa_{\ell, k}^{(i, j)} \geq 0$, $\forall i, \ell, k$.

end for

Stopping Criterion: $\sum_{j=1}^{N_V} \|\mathbf{x}_j^{(n+1)} - \mathbf{x}_j^{(n)}\|^2 < \delta$ for some $\delta > 0$.

Relaxation Parameters: Initialize $\lambda_j^{(0)} = \lambda_0$ and update using Algorithm 2 as

$$\lambda_j^{(n)} = \mathcal{J}(\tilde{\mathcal{F}}_j \cup \mathcal{F}_j^{(n)}, \lambda_j^{(n-1)}, \beta, \xi, \tilde{\mathbf{x}}_j^{(n)}) \quad (56)$$

for $j \in \{1, 2, \dots, N_V\}$, where $\tilde{\mathcal{F}}_j$ and \mathcal{F}_j are given by (S10) and (S11) in the supplementary material, respectively, and

$$\mathcal{S}_j^{(n)} \triangleq \{f \in \mathcal{F}_j \mid f(\tilde{\mathbf{x}}_j^{(n)}) \leq \gamma_{\ell, k}^{(i, j)}\}. \quad (57)$$

given by $\mathcal{O}\left(N_1 \sum_{k=1}^{K_j} |\tilde{S}_k^{(j)}|\right)$. Since the number of functions $\tilde{g}_{\ell, k}^{(j)}(\cdot)$ in (46) is equal to $\sum_{k=1}^{K_j} |\tilde{S}_k^{(j)}|$, evaluating (46) has the complexity $\mathcal{O}\left(\sum_{k=1}^{K_j} |\tilde{S}_k^{(j)}|\right)$. Similar to the case of noncooperative projections in (46), evaluating (47) requires a computational complexity of $\mathcal{O}\left(\sum_{k=1}^{K_j} \sum_{i=1, i \neq j}^{N_V} |S_k^{(i, j)}|\right)$. To analyze (51) and (52), assume that Algorithm 2 requires $\mathcal{O}(N_2)$ trials for determining a non-negative integer m . At each trial, (44) requires $\mathcal{O}(M)$ operations, where M is the number of functions at the input of Algorithm 2. Then, the computational complexity of step size selection in (51) and (52) is computed as $\mathcal{O}\left(N_2 \sum_{k=1}^{K_j} \left[|\tilde{S}_k^{(j)}| + \sum_{i=1, i \neq j}^{N_V} |S_k^{(i, j)}| \right]\right)$. Since the evaluation of the gradient projection operator in (39) requires $\mathcal{O}(1)$ operations, the averaging step in (50) has the complexity of $\mathcal{O}(1)$. Therefore, the overall complexity for the j th VLC unit is given by $\mathcal{O}\left(\sum_{k=1}^{K_j} \left[(N_1 + N_2) |\tilde{S}_k^{(j)}| + N_2 \sum_{i=1, i \neq j}^{N_V} |S_k^{(i, j)}| \right]\right)$.

Assuming that the number of iterations for the convergence of Algorithm 3 is on the order of $\mathcal{O}(N_3)$, the overall complexity of Algorithm 3 can be expressed as

$$\mathcal{O}\left(N_3 \sum_{j=1}^{N_V} \sum_{k=1}^{K_j} \left[\underbrace{(N_1 + N_2) |\tilde{S}_k^{(j)}|}_{\text{noncooperative}} + \underbrace{N_2 \sum_{i=1, i \neq j}^{N_V} |S_k^{(i, j)}|}_{\text{cooperative}} \right] \right) \quad (58)$$

where we can observe the contribution of the terms corresponding to noncooperative and cooperative projections separately.

2) *Complexity Analysis for Algorithm 4*: Following a similar approach to that in Algorithm 3, the complexity of Algorithm 4 can be obtained as

$$\mathcal{O}\left(N_3 \sum_{j=1}^{N_V} \sum_{k=1}^{K_j} \left[\underbrace{(N_1 + N_2) |\tilde{S}_k^{(j)}|}_{\text{noncooperative}} + N_2 \underbrace{\sum_{i=1, i \neq j}^{N_V} |S_k^{(i,j)}|}_{\text{cooperative}} \right] \right). \quad (59)$$

3) *Complexity Analysis for MLE in (9)*: Due to the nonconvexity of the MLE, we evaluate the complexity by assuming the use of the exhaustive search method over a bounded region. Suppose that each VLC unit can take N_{smp} different values in each of the three axes. Then, the number of possible locations for N_V VLC units is on the order of $\mathcal{O}(N_{\text{smp}}^{3N_V})$. At each search location, we need $\mathcal{O}\left(\sum_{j=1}^{N_V} \sum_{k=1}^{K_j} \left[|\tilde{S}_k^{(j)}| + \sum_{i=1, i \neq j}^{N_V} |S_k^{(i,j)}| \right] \right)$

operations for evaluating the cost function in (9). Hence, the overall complexity of the MLE can be computed as

$$\mathcal{O}\left(N_{\text{smp}}^{3N_V} \sum_{j=1}^{N_V} \sum_{k=1}^{K_j} \left[\underbrace{|\tilde{S}_k^{(j)}|}_{\text{noncooperative}} + \underbrace{\sum_{i=1, i \neq j}^{N_V} |S_k^{(i,j)}|}_{\text{cooperative}} \right] \right). \quad (60)$$

As observed from (58) and (59), the complexity of the proposed algorithms depends on the number of iterations and the size of the connectivity sets. On the other hand, the complexity of the MLE, expressed in (60), is exponential in the number of VLC units, which limits its scalability, as discussed in Section III-A. Hence, the proposed projection based algorithms provide low-complexity alternatives to the MLE in cooperative VLP scenarios.

V. CONVERGENCE ANALYSIS

In this section, the convergence analysis of the proposed algorithms in Algorithm 3 and Algorithm 4 is performed in the consistent case. To that aim, it is assumed that for each $j \in \{1, 2, \dots, N_V\}$, the intersection of the noncooperative and cooperative Lambertian sets in (24) is nonempty; that is, $\Lambda_j \cap \Upsilon_j \neq \emptyset$, where Λ_j and Υ_j are given by (25) and (26), respectively. In the following, we present the definitions of quasiconvexity and quasi-Fejér convergence, which will be used for the convergence proofs.

Definition 2 (*Quasiconvexity* [72]): A differentiable function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is quasiconvex if and only if $f(\mathbf{x}) \leq f(\mathbf{y})$ implies $\nabla f(\mathbf{y})^T (\mathbf{x} - \mathbf{y}) \leq 0 \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Definition 3 (*Quasi-Fejér Convergence* [38]): A sequence $\{\mathbf{y}^k\} \subset \mathbb{R}^n$ is quasi-Fejér convergent to a nonempty set V if for each $\mathbf{y} \in V$, there exists a non-negative integer M and a sequence $\{\epsilon^k\} \subset \mathbb{R}_{\geq 0}$ such that $\sum_{k=0}^{\infty} \epsilon_k < \infty$ and

$$\|\mathbf{y}^{k+1} - \mathbf{y}\|^2 \leq \|\mathbf{y}^k - \mathbf{y}\|^2 + \epsilon_k, \quad \forall k \geq M. \quad (61)$$

For the convergence analysis, we make the following assumptions:

- A1. Considering any $\mathbf{x}_j \in \Lambda_j \cap \Upsilon_j$ and $\hat{\mathbf{x}}_j \notin \Lambda_j \cap \Upsilon_j$, the inequality $g_{\ell,k}^{(i,j)}(\mathbf{x}_j, \mathbf{x}_i^{(n)}) \leq g_{\ell,k}^{(i,j)}(\hat{\mathbf{x}}_j, \mathbf{x}_i^{(n)})$ holds for every iteration index n and $\forall \ell, k, i, j$.
- A2. The sequence of path lengths taken by the iterations of the proposed algorithms are square summable, i.e.,

$$\sum_{n=0}^{\infty} \left(\|\tilde{\mathbf{x}}_j^{(n)} - \mathbf{x}_j^{(n)}\|^2 + \|\mathbf{x}_j^{(n+1)} - \tilde{\mathbf{x}}_j^{(n)}\|^2 \right) < \infty \text{ for } j \in \{1, 2, \dots, N_V\}.$$

Assumption A1 is valid especially when the cooperative algorithms can be initialized at some $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_{N_V})$ with $\mathbf{x}_j \in \Lambda_j, \forall j \in \{1, 2, \dots, N_V\}$. Assumption A1 implies that any point inside the intersection of the noncooperative and cooperative constraint sets is closer, in terms of the function value (whose zero-sublevel sets are the constraint sets), to the cooperative constraint sets than any point outside the intersection. When the iterations in the cooperative case start from coarse location estimates obtained in the absence of cooperation, the corresponding cooperative sets, which are dynamically changing at each iteration, may involve the set $\Lambda_j \cap \Upsilon_j$, but exclude the points outside $\Lambda_j \cap \Upsilon_j$, which yields $g_{\ell,k}^{(i,j)}(\mathbf{x}_j, \mathbf{x}_i^{(n)}) \leq 0 < g_{\ell,k}^{(i,j)}(\hat{\mathbf{x}}_j, \mathbf{x}_i^{(n)})$. On the other hand, Assumption A2 represents a realistic scenario through the Armijo rule in (43) and (44), which ensures a certain level of decline in the Lambertian functions at each iteration and generates a nonincreasing sequence of step sizes.

A. Quasi-Fejér Convergence

In the convergence analysis, the proof of convergence is based on the concept of quasi-Fejér convergent sequences, which possess nice properties that facilitate further investigation, as will be presented in Lemma 2. The following proposition establishes the quasi-Fejér convergence of the sequences generated by Algorithm 4 to the set $\Lambda_j \cap \Upsilon_j$.

Proposition 1: Assume A1 and A2 hold. Let $\{\mathbf{x}^{(n)}\}_{n=0}^{\infty}$ be any sequence generated by Algorithm 4, where $\mathbf{x}^{(n)} \triangleq (\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{N_V}^{(n)})$. Then, for each $j \in \{1, 2, \dots, N_V\}$, the sequence $\{\mathbf{x}_j^{(n)}\}_{n=0}^{\infty}$ is quasi-Fejér convergent to the set $\Lambda_j \cap \Upsilon_j$.

Proof: Please see Section S-II in the supplementary material.

The following proposition states the quasi-Fejér convergence of the sequences generated by Algorithm 3.

Proposition 2: Assume A1 and A2 hold. Let $\{\mathbf{x}^{(n)}\}_{n=0}^{\infty}$ be any sequence generated by Algorithm 3, where $\mathbf{x}^{(n)} \triangleq (\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{N_V}^{(n)})$. Then, for each $j \in \{1, 2, \dots, N_V\}$, the sequence $\{\mathbf{x}_j^{(n)}\}_{n=0}^{\infty}$ is quasi-Fejér convergent to the set $\Lambda_j \cap \Upsilon_j$.

Proof: Please see Section S-III in the supplementary material.

As the quasi-Fejér convergence of the sequences generated by the proposed algorithms is stated, the following lemma presents the properties of quasi-Fejér convergent sequences.

Lemma 2 (*Theorem 4.1 in [38]*): If a sequence $\{\mathbf{y}^k\}$ is quasi-Fejér convergent to a nonempty set V , the following conditions hold:

- 1) $\{\mathbf{y}^k\}$ is bounded.
- 2) If V contains an accumulation point of $\{\mathbf{y}^k\}$, then $\{\mathbf{y}^k\}$ converges to a point $\mathbf{y} \in V$.

B. Limiting Behavior of Step Size Sequences

In this part, we investigate the limiting behavior of the step size sequences, which are updated according to the procedure in Algorithm 2. The following two lemmas prove that the step size sequences generated by Algorithm 4 and Algorithm 3 have positive limits.

Lemma 3: Any step size sequence $\lambda_j^{(n)}$ generated by Algorithm 4 has a positive limit, i.e.,

$$\lim_{n \rightarrow \infty} \lambda_j^{(n)} > 0. \quad (62)$$

Proof: Please see Section S-IV in the supplementary material.

Lemma 4: Any step size sequences $\lambda_{j,nc}^{(n)}$ and $\lambda_{j,c}^{(n)}$ generated by Algorithm 3 have positive limits, i.e.,

$$\lim_{n \rightarrow \infty} \lambda_{j,nc}^{(n)} > 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \lambda_{j,c}^{(n)} > 0. \quad (63)$$

Proof: Please see Section S-V in the supplementary material.

Lemma 3 and Lemma 4 will prove to be useful for deriving the fundamental convergence properties of the proposed algorithms, as investigated next.

C. Main Convergence Results

In this part, we present the main convergence results for the proposed algorithms, i.e., convergence to a solution of Problem 1.

Proposition 3: Let $\{\mathbf{x}^{(n)}\}_{n=0}^{\infty}$ be any sequence generated by Algorithm 4, where $\mathbf{x}^{(n)} \triangleq (\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{N_V}^{(n)})$. Then, for each $j \in \{1, 2, \dots, N_V\}$, the sequence $\{\mathbf{x}_j^{(n)}\}_{n=0}^{\infty}$ converges to a point $\mathbf{x}_j \in \Lambda_j \cap \Upsilon_j$, i.e., a solution of Problem 1.

Proof: Please see Section S-VI in the supplementary material. ■

Proposition 4: Let $\{\mathbf{x}^{(n)}\}_{n=0}^{\infty}$ be any sequence generated by Algorithm 3, where $\mathbf{x}^{(n)} \triangleq (\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{N_V}^{(n)})$. Then, for each $j \in \{1, 2, \dots, N_V\}$, the sequence $\{\mathbf{x}_j^{(n)}\}_{n=0}^{\infty}$ converges to a point $\mathbf{x}_j \in \Lambda_j \cap \Upsilon_j$, i.e., a solution of Problem 1.

Proof: Please see Section S-VII in the supplementary material. ■

VI. NUMERICAL RESULTS

In this section, numerical examples are provided to investigate the theoretical bounds on cooperative localization in VLP networks and to evaluate the performance of the proposed projection-based algorithms. The VLP network parameters are determined in a similar manner to the work in [12] and [13]. The area of each PD is set to 1 cm^2 and the Lambertian order of all the LEDs is selected as $m = 1$. In addition, the noise variances are calculated using [73, Eq. 6]. The parameters for noise variance calculation are set to be the same as those used in [73] (see Table I in [73]).

The VLP network considered in the simulations is illustrated in Fig. 3. A room of size $10\text{m} \times 10\text{m} \times 5\text{m}$ is considered, where there exist $N_L = 4$ LED transmitters on the ceiling which are located at $\mathbf{y}_1 = [1 \ 1 \ 5]^T \text{m}$, $\mathbf{y}_2 = [1 \ 9 \ 5]^T \text{m}$, $\mathbf{y}_3 = [9 \ 1 \ 5]^T \text{m}$, and $\mathbf{y}_4 = [9 \ 9 \ 5]^T \text{m}$. The LEDs on the ceiling have perpendicular orientations, i.e., $\tilde{\mathbf{n}}_{T,j} = [0 \ 0 \ -1]^T$ for $j \in \{1, 2, 3, 4\}$. In addition, there exist $N_V = 2$ VLC units whose locations are given by $\mathbf{x}_1 = [2 \ 5 \ 1]^T \text{m}$ and $\mathbf{x}_2 = [6 \ 6 \ 1.5]^T \text{m}$. Each VLC unit consists of two PDs and one LED, with offsets with respect to the center of the VLC unit being set to $\mathbf{a}_{j,1} = [0 \ -0.1 \ 0]^T \text{m}$, $\mathbf{a}_{j,2} = [0 \ 0.1 \ 0]^T \text{m}$, and $\mathbf{b}_{j,1} = [0.1 \ 0 \ 0]^T \text{m}$ for $j = 1, 2$. The orientation vectors of the PDs and the LEDs on the VLC units are obtained as the normalized versions (the orientation vectors are unit-norm) of the following vectors: $\mathbf{n}_{R,1}^{(1)} = [0.3 \ -0.1 \ 1]^T$, $\mathbf{n}_{R,1}^{(2)} = [0.2 \ 0.4 \ 1]^T$, $\mathbf{n}_{R,2}^{(1)} = [0.8 \ 0.6 \ 0.1]^T$, $\mathbf{n}_{R,2}^{(2)} = [-0.7 \ 0.2 \ 0.1]^T$, $\mathbf{n}_{T,1}^{(1)} = [0.9 \ 0.4 \ 0.1]^T$, and $\mathbf{n}_{T,1}^{(2)} = [-0.8 \ 0.1 \ 0.1]^T$. Furthermore, the connectivity sets are defined as $S_1^{(i,j)} = \emptyset$, $S_2^{(i,j)} = \{1\}$ for $i, j \in \{1, 2\}$, $i \neq j$ for the cooperative measurements and $\tilde{S}_1^{(1)} = \{1, 2, 3\}$, $\tilde{S}_1^{(2)} = \{2, 3, 4\}$ and $\tilde{S}_2^{(j)} = \emptyset$ for $j \in \{1, 2\}$ for the noncooperative measurements.

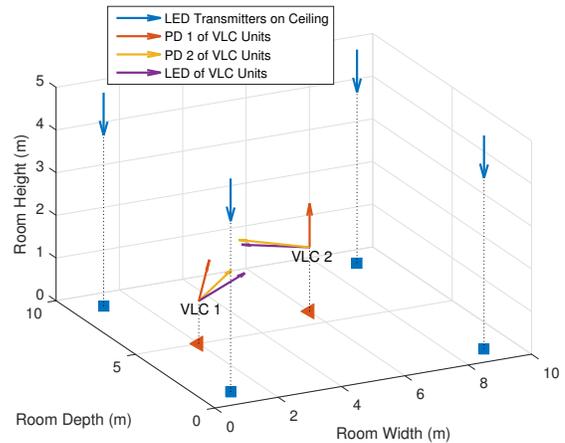


Fig. 3. VLP network configuration in the simulations. Each VLC unit contains two PDs and one LED. PD 1 of the VLC units is used to obtain measurements from the LEDs on the ceiling while PD 2 of the VLC units communicates with the LED of the other VLC unit for cooperative localization. The squares and the triangles show the projections of the LEDs and the VLC units on the floor, respectively.

A. Theoretical Bounds

In this part, the CRLB expression derived in Section II is investigated to illustrate the effects of cooperation on the localization performance of VLP networks.

1) *Performance with Respect to Transmit Power of LEDs on Ceiling:* In order to analyze the localization performance of the VLC units with respect to the transmit powers of the LEDs on the ceiling (equivalently, anchors), individual CRLBs for localization of the VLC units in noncooperative and cooperative scenarios are plotted against the transmit powers of LEDs on the ceiling in Fig. 4, where the transmit powers of the VLC units are fixed to 1 W. As observed from Fig. 4, cooperation among VLC units can provide substantial improvements in localization accuracy (about 43 cm and 14 cm improvement, respectively, for VLC 1 and VLC 2 for the LED transmit power of 300 mW). We note that the improvement gained by employing cooperation is higher for VLC 1 as compared to that for VLC 2. This is an intuitive result since the localization of VLC 1 depends mostly on LED 1 and LED 2 (the other LEDs are not sufficiently close to facilitate the localization process), and incorporating cooperative measurements for VLC 1 provides an enhancement in localization performance that is much greater than that for VLC 2, which can obtain informative measurements from the LEDs on the ceiling even in the absence of cooperation as seen from the network geometry in Fig. 3. In addition, the CRLBs in the cooperative scenario converge to those in the noncooperative scenario as the transmit powers of the LEDs increase. Since the first (second) summand in the FIM expression in (12) corresponds to the noncooperative (cooperative) localization, higher transmit powers of the LEDs on the ceiling cause the first summand to be much greater than the second one, which makes the contribution of cooperation to the FIM negligible. Hence, the effect of cooperation on localization performance becomes more significant as the transmit power decreases, which is in compliance with the results obtained for RF based cooperative localization networks [20].

2) *Performance with Respect to Transmit Power of VLC Units:* Secondly, the localization performance of the VLC units is investigated with respect to the transmit powers of the VLC units when the transmit powers of the LEDs on the ceiling are fixed to 1 W. Fig. 5 illustrates the CRLBs for localization of the VLC units against the transmit powers of the VLC units in the noncooperative and cooperative cases. As observed from Fig. 5, cooperation leads to a higher improvement in the performance of VLC 1, similar to Fig. 4. In addition, via the

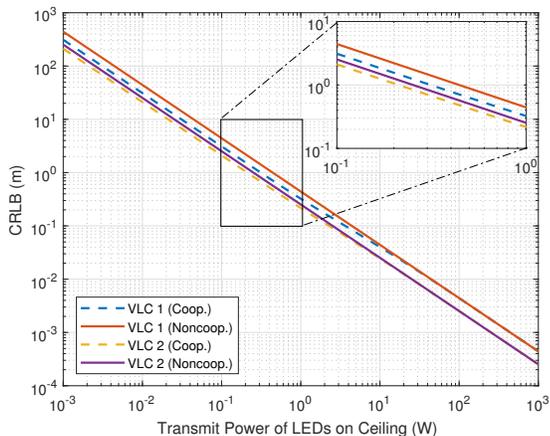


Fig. 4. Individual CRLBs for localization of VLC units in both noncooperative and cooperative cases with respect to the transmit power of LEDs on ceiling, where

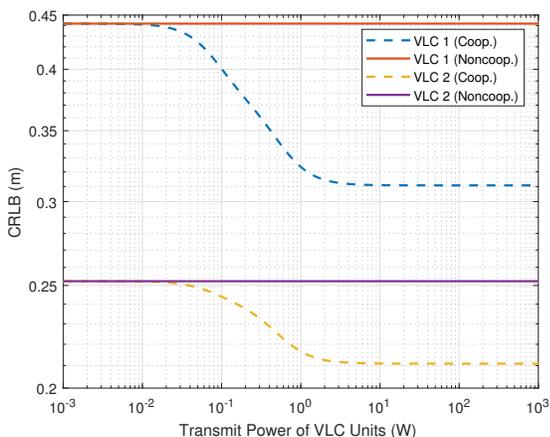


Fig. 5. Individual CRLBs for localization of VLC units in both noncooperative and cooperative cases with respect to the transmit power of VLC units, where the transmit power of LEDs on ceiling is taken as 1W.

FIM expression in (12), it can be noted that the contribution of cooperation to localization performance gets higher as the transmit powers of the VLC units increase, which is also observed from Fig. 5. However, the CRLB reaches a saturation level above a certain power threshold, as opposed to Fig. 4, where the CRLB continues to decrease as the power increases. The main reason for this distinction between the effects of the transmit powers of the LEDs on the ceiling and those of the VLC units can be explained as follows: For a fixed transmit power of the VLC units, the localization error by using three anchors (i.e., three LEDs on the ceiling that are connected to the corresponding VLC unit) converges to zero as the transmit powers of the anchors increase regardless of the existence of cooperation. On the other hand, for a fixed transmit power of the LEDs on the ceiling, increasing the transmit power of the VLC unit (i.e., one of the anchors) cannot reduce the localization error below a certain level. Therefore, the saturation level represents the localization accuracy that can be attained by four anchors with three anchors leading to noisy RSS measurements and one anchor generating noise-free RSS measurements.

B. Performance of the Proposed Algorithms

In this part, the proposed algorithms in Algorithm 3 (CCGP) and Algorithm 4 (CSGP) are evaluated in terms of localization performance and convergence speed. For both algorithms, the initial step size is selected as $\lambda_0 = 1$, the step size shrinkage factor and the degree of decline in the Armijo rule in Algorithm 2 are set to $\xi = 0.5$ and $\beta = 0.001$, respectively. The VLC units are initialized at the positions of the closest LEDs

on the ceiling which are connected to the corresponding VLC units.

Localization performances of the algorithms are presented in both the absence and the presence of cooperation and compared against those of the ML estimator in (9) and the CRLBs derived in Section II. In order to ensure convergence to the global minimum, the ML estimator is implemented using a multi-start optimization algorithm with 100 initial points randomly selected from the interval $[0 \ 10]$ m at each axis.¹⁰ In addition, two different measurement noise distributions, namely, Gaussian and exponential, are considered while evaluating the proposed algorithms as in [28]. The Gaussian noise is used to model the case in which the RSS measurement noise can be both positive and negative, whereas the exponentially distributed noise (subtracted from the true value) represents the scenario in which the RSS measurements are negatively biased, which leads to the feasibility modeling of the localization problem in Section III-B. Furthermore, the average residuals at each iteration are calculated to assess the convergence speed of the proposed algorithms [29]:

$$\varrho_n = \frac{1}{MN_V} \sum_{m=1}^M \|\mathbf{x}^{(n,m)} - \mathbf{x}^{(n-1,m)}\| \quad (64)$$

where $\mathbf{x}^{(n,m)} = (\mathbf{x}_1^{(n,m)}, \dots, \mathbf{x}_{N_V}^{(n,m)})$ denotes the position vector of all the VLC units at the n th iteration for the m th Monte Carlo realization of measurement noises and M is the number of Monte Carlo realizations.

In the simulations, two-dimensional localization is performed by assuming that the VLC units have known heights. Therefore, with the knowledge of perpendicular LED orientations, Case 2 type Lambertian sets in Section III-D2 are utilized for localization based on the measurements from the LEDs on the ceiling. The cooperation among the VLC units is modeled by Case 1 type Lambertian sets in Section III-D1.

1) *Gaussian Noise*: In Fig. 6, the average localization errors of the VLC units for the different algorithms are plotted against the transmit power of the LEDs on the ceiling for the case of the Gaussian measurement noise by fixing the transmit powers of the LEDs at the VLC units to 1W. From Fig. 6, it is observed that the cooperative approach can significantly reduce the localization errors, especially in the low SNR regime (about 60 cm and 70 cm reduction for CSGP and CCGP algorithms, respectively, for 100 mW LED transmit power). In addition, both Algorithm 3 (CCGP) and Algorithm 4 (CSGP) can attain the localization error levels that asymptotically converge to zero at the same rate as that of the CRLB. Moreover, it can be inferred from Fig. 6 that the proposed iterative methods achieve higher localization performance than the ML estimator in the low SNR regime for both the noncooperative and the cooperative scenarios. Although the ML estimator is forced to converge to the global minimum via the multi-start optimization procedure involving 100 different executions of a local solver, whose complexity may be prohibitive for practical implementations, it has lower performance than the proposed approaches, which depend on low-complexity iterative gradient projections. Hence, at low SNRs, the proposed algorithms are superior to the MLE in terms of both the localization performance and the computational complexity. Furthermore, the simultaneous projections outperforms the cyclic projections at low SNRs at the cost of a higher number of set projections, but the two approaches converge asymptotically as the SNR increases.

¹⁰The implemented estimator is effectively a maximum *a posteriori* probability (MAP) estimator with a uniform prior distribution over the interval $[0 \ 10]$ m, based on the prior information that VLC units are inside the room. Hence, the implemented ML estimator may achieve smaller RMSEs than the CRLB in the low SNR regime, where the prior information becomes more significant as the measurements are very noisy.

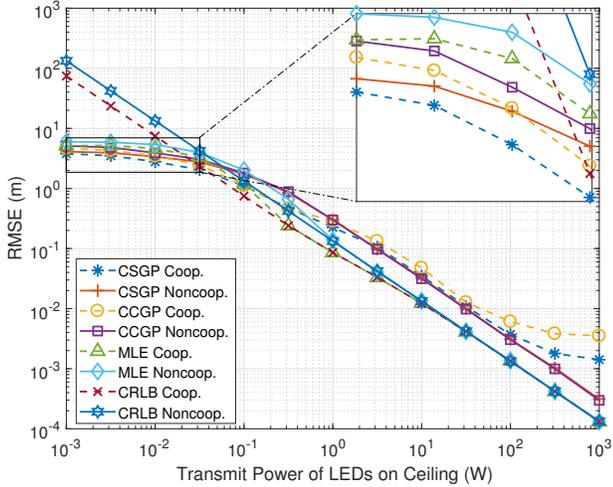
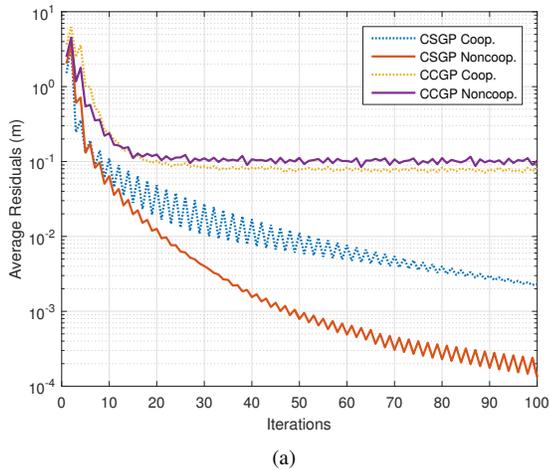
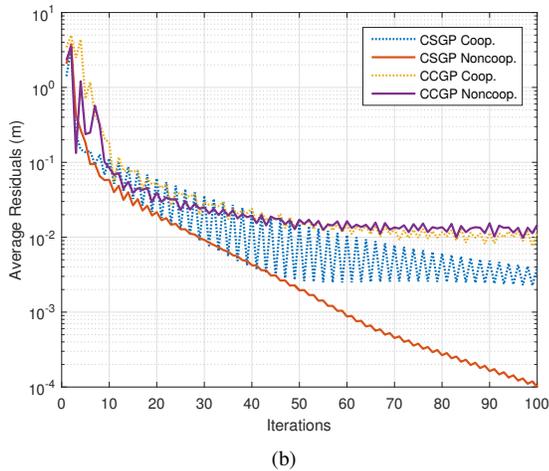


Fig. 6. Average localization error of VLC units with respect to the transmit power of LEDs on ceiling for the proposed algorithms in Algorithm 3 (CCGP) and Algorithm 4 (CSGP) along with the MLE and CRLB for the case of Gaussian measurement noise.



(a)



(b)

Fig. 7. Convergence rate of the average residuals in (64) for the proposed algorithms in Algorithm 3 and Algorithm 4 for the case of Gaussian measurement noise, where the transmit power of LEDs on ceiling is (a) 100 mW and (b) 1 W.

Fig. 7(a) and Fig. 7(b) report the average residuals calculated by (64) corresponding to the proposed algorithms versus the number of iterations for 100 mW and 1 W of transmit powers of the LEDs on the ceiling, respectively. CSGP in the absence of cooperation has the fastest convergence rate and exhibits an almost monotonic convergence behavior. However, CSGP in the cooperative scenario shows relatively slow convergence in general and a locally nonmonotonic behavior when several consecutive iterations are taken into account. This is due to the cooperative Lambertian sets being involved in the simultaneous projection operations. In addition, cyclic projections tend to settle into limit cycle oscillations¹¹ after few iterations, thus implying that the sequence itself does not converge to a point, but it has several subsequences that converge [44]. This behavior, called *cyclic convergence*, is encountered in cyclic (sequential) projections if the feasibility problem is inconsistent [44], [74]. Furthermore, by comparing Fig. 7(a) and Fig. 7(b), it is observed that the magnitude of limit cycle oscillations in the CCGP algorithm gets smaller as the SNR increases since the region of uncertainty becomes narrower at higher SNR values, thereby making the convergent subsequences close to each other.

2) *Exponential Noise*: To investigate the performance of the algorithms under exponentially distributed measurement noise, the average localization errors are plotted against the transmit power of the LEDs on the ceiling for the case of the subtractive exponential noise in Fig. 8. Similar to the case of the Gaussian noise, the proposed algorithms succeed in converging to the true VLC unit positions as the SNR increases. Since the projection based methods rely on the assumption of negatively biased measurements, they perform slightly better at low SNRs as compared to the case of the Gaussian noise. On the other hand, the MLE produces larger errors at low SNRs for the exponentially distributed noise since its derivation is based on the assumption of Gaussian noise.

The average residuals in the case of the exponentially distributed noise are illustrated in Fig. 9(a) and Fig. 9(b) for two different LED power levels. In contrary to the case of Gaussian noise, cyclic projections do not fall into limit cycles and provide globally monotonic convergence results as the feasibility problem is consistent, which complies with the results presented in the literature pertaining to the study of CFPs [44]. In addition, it is observed that both the cyclic and the sequential projection methods have faster convergence for lower SNR values since it takes fewer iterations to get inside the intersection of the constraint sets, which becomes larger as the SNR decreases.¹²

VII. CONCLUDING REMARKS

In this paper, a cooperative VLP network has been proposed based on a generic system model consisting of LED transmitters at known locations and VLC units with multiple LEDs and PDs. First, the CRLB on the overall localization error of the VLC units has been derived to quantify the effects of cooperation on the localization accuracy of VLP networks. Then, due to the nonconvex nature of the corresponding ML expression, the problem of cooperative localization has been formulated as

¹¹ In inconsistent feasibility problems (i.e., those with empty region of intersection), the sequence of points obtained by cyclic projections does not converge [44]. However, it has convergent subsequences each of which converges to a different point [44]. In the limit, the cyclic projections based algorithm (i.e., CCGP) visits each of these limiting points in a sequential fashion, leading to a cyclic behavior called *limit cycle*. In this case, average residuals in (64) (i.e., distances between consecutive iteration points) can be regarded as *limit cycle oscillations*.

¹² Since the measured RSS is always smaller than the true RSS value in the case of negative exponential noise (see (16)), the Lambertian sets become larger at lower SNR values (as the size of a Lambertian set \mathcal{L} in (19) is inversely proportional to the corresponding RSS measurement \hat{P}_r , or, equivalently γ in (18)).

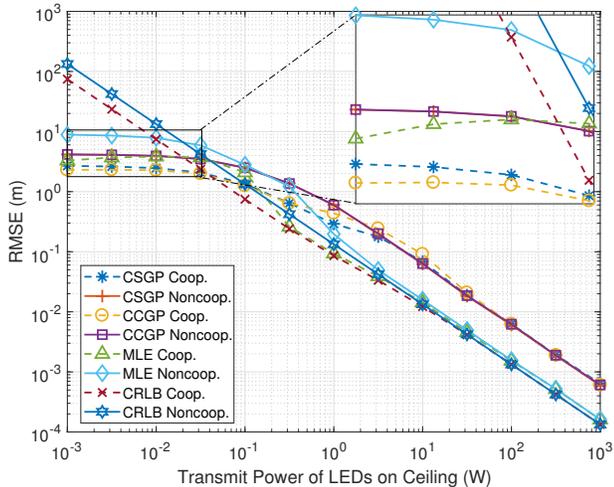


Fig. 8. Average localization error of VLC units with respect to the transmit power of LEDs on ceiling for the proposed algorithms in Algorithm 3 (CCGP) and Algorithm 4 (CSGP) along with the MLE and CRLB for the case of exponentially distributed measurement noise.

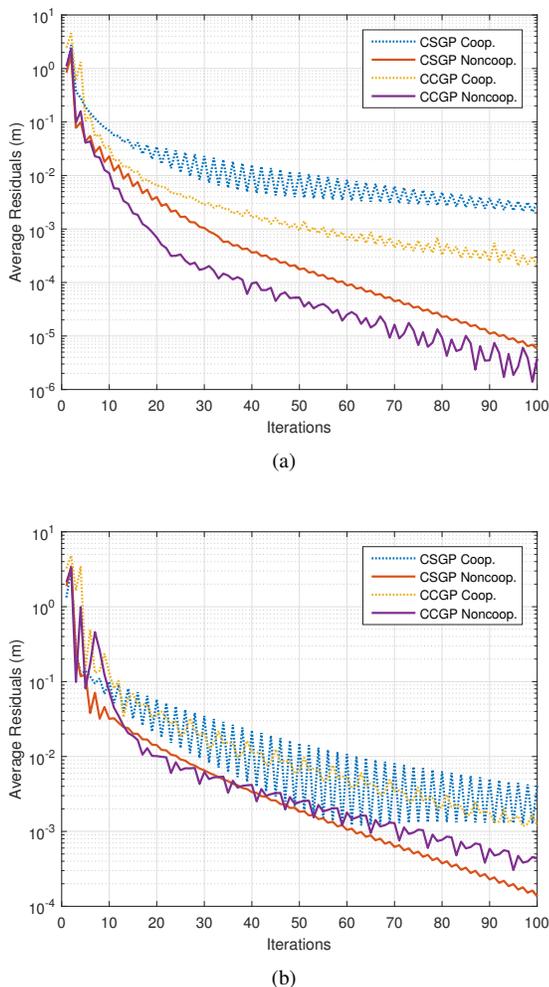


Fig. 9. Convergence rate of the average residuals in (64) for the proposed algorithms in Algorithm 3 and Algorithm 4 for the case of exponentially distributed measurement noise, where the transmit power of LEDs on ceiling is (a) 100 mW and (b) 1 W.

a QFP, which facilitates the development of low-complexity decentralized feasibility-seeking methods. In order to solve the feasibility problem, iterative gradient projections based algorithms have been proposed. Furthermore, based on the notion of quasi-Fejér convergent sequences, formal convergence proofs have been provided for the proposed algorithms in the consistent case. Finally, numerical examples have been presented to illustrate the significance of cooperation in VLP networks and to investigate the performance of the proposed algorithms in terms of localization accuracy and convergence speed. It has been verified that the proposed iterative methods asymptotically converge to the true positions of VLC units at high SNR and exhibit superior performance over the ML estimator at low SNRs in terms of both implementation complexity and localization accuracy.

An important research direction for future studies is to explore the convergence properties of Algorithm 3 and Algorithm 4 when the proposed QFP is inconsistent. In the inconsistent case, simultaneous projection algorithms tend to converge to a minimizer of a *proximity function* that specifies the distance to constraint sets [44], [75]. For the implicit CFP (ICFP) considered in TOA-based wireless network localization, the POCS based simultaneous algorithm is shown to converge to the minimizer of a convex function, which is the sum of squares of the distances to the constraint sets [29]. Therefore, finding proximity functions characterizing the behavior of simultaneous projections (e.g., Algorithm 4) for the inconsistent QFPs [37] would be a significant extension for the set-theoretic estimation literature.

REFERENCES

- [1] G. Mao and B. Fidan, *Localization Algorithms and Strategies for Wireless Sensor Networks*. Information Science Reference, 2009.
- [2] K. Pahlavan and P. Krishnamurthy, *Principles of Wireless Access and Localization*, ser. Wiley Desktop Editions. Wiley, 2013.
- [3] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.
- [4] S. Gezici, "A survey on wireless position estimation," *Wireless Personal Communications*, vol. 44, no. 3, pp. 263–282, Feb. 2008.
- [5] R. Zekavat and R. M. Buehrer, *Handbook of Position Location: Theory, Practice and Advances*. John Wiley & Sons, 2011.
- [6] Z. Sahinoglu, S. Gezici, and I. Guvenc, *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols*. New York: Cambridge University Press, 2008.
- [7] J. Armstrong, Y. Sekercioglu, and A. Neild, "Visible light positioning: A roadmap for international standardization," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 68–73, Dec. 2013.
- [8] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: A survey, potential and challenges," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2047–2077, Fourthquarter 2015.
- [9] H. Burchardt, N. Serafimovski, D. Tsonev, S. Videv, and H. Haas, "VLC: Beyond point-to-point communication," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 98–105, July 2014.
- [10] X. Zhang, J. Duan, Y. Fu, and A. Shi, "Theoretical accuracy analysis of indoor visible light communication positioning system based on received signal strength indicator," *Journal of Lightwave Technology*, vol. 32, no. 21, pp. 4180–4186, Nov. 2014.
- [11] W. Zhang, M. I. S. Chowdhury, and M. Kavehrad, "Asynchronous indoor positioning system based on visible light communications," *Optical Engineering*, vol. 53, no. 4, pp. 045 105–1–045 105–9, 2014.
- [12] T. Wang, Y. Sekercioglu, A. Neild, and J. Armstrong, "Position accuracy of time-of-arrival based ranging using visible light with application in indoor localization systems," *Journal of Lightwave Technology*, vol. 31, no. 20, pp. 3302–3308, Oct. 2013.
- [13] M. F. Keskin and S. Gezici, "Comparative theoretical analysis of distance estimation in visible light positioning systems," *Journal of Lightwave Technology*, vol. 34, no. 3, pp. 854–865, Feb. 2016.
- [14] S.-Y. Jung, S. Hann, and C.-S. Park, "TDOA-based optical wireless indoor localization using LED ceiling lamps," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 4, pp. 1592–1597, Nov. 2011.
- [15] Y. Eroglu, I. Guvenc, N. Pala, and M. Yuksel, "AOA-based localization and tracking in multi-element VLC systems," in *IEEE 16th Annual Wireless and Microwave Technology Conference (WAMICON)*, Apr. 2015.
- [16] H. Steendam, T. Q. Wang, and J. Armstrong, "Theoretical lower bound for indoor visible light positioning using received signal strength measurements and an aperture-based receiver," *Journal of Lightwave Technology*, vol. 35, no. 2, pp. 309–319, Jan 2017.

- [17] M. F. Keskin, E. Gonendik, and S. Gezici, "Improved lower bounds for ranging in synchronous visible light positioning systems," *Journal of Lightwave Technology*, vol. 34, no. 23, pp. 5496–5504, Dec 2016.
- [18] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54–69, July 2005.
- [19] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 427–450, Feb. 2009.
- [20] M. R. Gholami, M. F. Keskin, S. Gezici, and M. Jansson, *Cooperative Positioning in Wireless Networks*. Wiley Encyclopedia of Electrical and Electronics Engineering, 2016, pp. 1–19.
- [21] P. Biswas, T.-C. Lian, T.-C. Wang, and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Transactions on Sensor Networks*, vol. 2, no. 2, pp. 188–220, 2006.
- [22] K. W. K. Lui, W. K. Ma, H. C. So, and F. K. W. Chan, "Semidefinite programming algorithms for sensor network node localization with uncertainties in anchor positions and/or propagation speed," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 752–763, Feb 2009.
- [23] R. W. Ouyang, A. K. S. Wong, and C. T. Lea, "Received signal strength-based wireless localization via semidefinite programming: Noncooperative and cooperative schemes," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1307–1318, March 2010.
- [24] P. Tseng, "Second-order cone programming relaxation of sensor network localization," *SIAM Journal on Optimization*, vol. 18, no. 1, pp. 156–185, Feb. 2007.
- [25] G. Naddafzadeh-Shirazi, M. B. Shenouda, and L. Lampe, "Second order cone programming for sensor network localization with anchor position uncertainty," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 749–763, February 2014.
- [26] C. Soares, J. Xavier, and J. Gomes, "Simple and fast convex relaxation method for cooperative localization in sensor networks using range measurements," *IEEE Transactions on Signal Processing*, vol. 63, no. 17, pp. 4532–4543, Sept 2015.
- [27] D. Blatt and A. O. Hero, "Energy-based sensor network source localization via projection onto convex sets," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3614–3619, Sept 2006.
- [28] M. R. Gholami, H. Wymeersch, E. G. Ström, and M. Rydström, "Wireless network positioning as a convex feasibility problem," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 1, p. 161, 2011.
- [29] M. R. Gholami, L. Tetrushvili, E. G. Ström, and Y. Censor, "Cooperative wireless sensor network positioning via implicit convex feasibility," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5830–5840, Dec 2013.
- [30] Y. Zhang, Y. Lou, Y. Hong, and L. Xie, "Distributed projection-based algorithms for source localization in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3131–3142, June 2015.
- [31] J. A. Costa, N. Patwari, and A. O. Hero, III, "Distributed weighted-multidimensional scaling for node localization in sensor networks," *ACM Transactions on Sensor Networks*, vol. 2, no. 1, pp. 39–64, Feb. 2006.
- [32] M. Kavehrad, "Sustainable energy-efficient wireless applications using light," *IEEE Communications Magazine*, vol. 48, no. 12, pp. 66–73, 2010.
- [33] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "LED based indoor visible light communications: State of the art," *IEEE Communications Surveys Tutorials*, vol. 17, no. 3, pp. 1649–1678, thirdquarter 2015.
- [34] D. Tagliaferri and C. Capsoni, "High-speed wireless infrared uplink scheme for airplane passengers communications," *Electronics Letters*, vol. 53, no. 13, pp. 887–888, 2017.
- [35] C. Quintana, V. Guerra, J. Rufo, J. Rabadan, and R. Perez-Jimenez, "Reading lamp-based visible light communication system for in-flight entertainment," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 1, pp. 31–37, 2013.
- [36] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, 1997.
- [37] Y. Censor and A. Segal, "Algorithms for the quasiconvex feasibility problem," *Journal of Computational and Applied Mathematics*, vol. 185, no. 1, pp. 34 – 50, 2006.
- [38] A. N. Iusem, B. F. Svaiter, and M. Teboulle, "Entropy-like proximal methods in convex programming," *Mathematics of Operations Research*, vol. 19, no. 4, pp. 790–814, 1994.
- [39] T. Jia and R. M. Buehrer, "A set-theoretic approach to collaborative position location for wireless networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 9, pp. 1264–1275, 2011.
- [40] A. Carmi, Y. Censor, and P. Gurfil, "Convex feasibility modeling and projection methods for sparse signal recovery," *Journal of Computational and Applied Mathematics*, vol. 236, no. 17, pp. 4318–4335, 2012.
- [41] P. L. Combettes, "Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections," *IEEE Transactions on Image Processing*, vol. 6, no. 4, pp. 493–506, Apr 1997.
- [42] Y. Censor, A. Gibali, F. Lenzen, and C. Schnorr, "The implicit convex feasibility problem and its application to adaptive image denoising," *arXiv:1606.05848*, 2016.
- [43] Y. Censor and A. Segal, "Iterative projection methods in biomedical inverse problems," in *Proceeding of the interdisciplinary workshop on Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT)*, Pisa, Italy, Oct. 2008, pp. 65–96.
- [44] D. Butnariu, Y. Censor, P. Gurfil, and E. Hadar, "On the behavior of subgradient projections methods for convex feasibility problems in Euclidean spaces," *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 786–807, 2008.
- [45] A. Sahin, Y. S. Eroglu, I. Guvenc, N. Pala, and M. Yuksel, "Hybrid 3-D localization for visible light communication systems," *Journal of Lightwave Technology*, vol. 33, no. 22, pp. 4589–4599, Nov 2015.
- [46] M. Yasir, S. W. Ho, and B. N. Vellambi, "Indoor positioning system using visible light and accelerometer," *Journal of Lightwave Technology*, vol. 32, no. 19, pp. 3306–3316, Oct. 2014.
- [47] Y. E. Ustev, O. Durmaz Incel, and C. Ersoy, "User, device and orientation independent human activity recognition on mobile phones: Challenges and a proposal," in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 2013, pp. 1427–1436.
- [48] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A reliable and accurate indoor localization method using phone inertial sensors," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 421–430.
- [49] V. Renaudin and C. Combettes, "Magnetic, acceleration fields and gyroscope quaternion (MAGYQ)-based attitude estimation with smartphone sensors for indoor pedestrian navigation," *Sensors*, vol. 14, no. 12, pp. 22 864–22 890, 2014.
- [50] L. Li, P. Hu, C. Peng, G. Shen, and F. Zhao, "Epsilon: A visible light based positioning system," in *11th USENIX Symposium on Networked Systems Design and Implementation*, Seattle, WA, 2014, pp. 331–343.
- [51] Z. Ghassenlooy, L. N. Alves, S. Zvanovec, and M.-A. Khalighi, *Visible Light Communications: Theory and Applications*. CRC Press, 2017.
- [52] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [53] "Eye safety risk assessment of infrared emitting diodes according IEC 62471 (based on CIE s009)." [Online]. Available: <https://www.vishay.com/docs/81935/eyesafe.pdf>
- [54] S. Penfold, R. Zalas, M. Casiraghi, M. Brooke, Y. Censor, and R. Schulte, "Sparsity constrained split feasibility for dose-volume constraints in inverse planning of intensity-modulated photon or proton therapy," *Physics in Medicine and Biology*, vol. 62, p. 3599, May 2017.
- [55] S. Yousefi, H. Wymeersch, X. W. Chang, and B. Champagne, "Tight two-dimensional outer-approximations of feasible sets in wireless sensor networks," *IEEE Communications Letters*, vol. 20, no. 3, pp. 570–573, March 2016.
- [56] M. R. Gholami, M. Rydström, and E. G. Ström, "Positioning of node using plane projection onto convex sets," in *2010 IEEE Wireless Communication and Networking Conference*, April 2010, pp. 1–5.
- [57] M. R. Gholami, S. Gezici, E. G. Ström, and M. Rydström, "A distributed positioning algorithm for cooperative active and passive sensors," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2010.
- [58] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Review*, vol. 38, no. 3, pp. 367–426, 1996.
- [59] Y. Mathlouthi, A. Mitiche, and I. B. Ayed, *Boundary Preserving Variational Image Differentiation*. Cham: Springer International Publishing, 2016, pp. 355–364.
- [60] M. Bazarra, H. Sherali, and C. Shetty, *Nonlinear Programming: Theory and Algorithms*. Wiley-Interscience, 2006.
- [61] S.-H. Yang, E.-M. Jung, and S.-K. Han, "Indoor location estimation based on LED visible light communication using multiple optical receivers," *IEEE Communications Letters*, vol. 17, no. 9, pp. 1834–1837, Sep. 2013.
- [62] Y. Censor and A. Lent, "Cyclic subgradient projections," *Mathematical Programming*, vol. 24, no. 1, pp. 233–235, 1982.
- [63] P. L. Combettes, "Quasi-Fejérian analysis of some optimization algorithms," in *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, ser. Studies in Computational Mathematics, Y. C. Dan Butnariu and S. Reich, Eds. Elsevier, 2001, vol. 8, pp. 115 – 152.
- [64] J. Von Neumann, *Functional Operators (AM-22), Volume 2: The Geometry of Orthogonal Spaces.(AM-22)*. Princeton University Press, 2016, vol. 2.
- [65] I. Halperin, "The product of projection operators," *Acta Sci. Math.(Szeged)*, vol. 23, no. 1, pp. 96–99, 1962.
- [66] L. Armijo, "Minimization of functions having Lipschitz continuous first partial derivatives," *Pacific J. Math.*, vol. 16, no. 1, pp. 1–3, 1966.
- [67] P. Wolfe, "Convergence conditions for ascent methods," *SIAM Review*, vol. 11, no. 2, pp. 226–235, 1969.
- [68] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.
- [69] J. Fliege and B. F. Svaiter, "Steepest descent methods for multicriteria optimization," *Mathematical Methods of Operations Research*, vol. 51, no. 3, pp. 479–494, 2000.
- [70] L. T. D. Santos, "A parallel subgradient projections method for the convex feasibility problem," *Journal of Computational and Applied Mathematics*, vol. 18, no. 3, pp. 307 – 320, 1987.
- [71] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE/ACM Transactions on Networking (TON)*, vol. 14, no. SI, pp. 2508–2530, 2006.
- [72] O. L. Mangasarian, *Nonlinear programming*. SIAM, 1994.
- [73] H. Ma, L. Lampe, and S. Hranilovic, "Coordinated broadcasting for multiuser indoor visible light communication systems," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3313–3324, Sep. 2015.
- [74] L. Gubin, B. Polyak, and E. Raik, "The method of projections for finding the common point of convex sets," *{USSR} Computational Mathematics and Mathematical Physics*, vol. 7, no. 6, pp. 1 – 24, 1967.

- [75] Y. Censor, A. R. D. Pierro, and M. Zaknoon, "Steered sequential projections for the inconsistent convex feasibility problem." *Nonlinear analysis: theory, method, and application, series A*, vol. 59, pp. 385–405, 2004.



Musa Furkan Keskin received the B.S. and M.S. degrees from the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey, in 2010 and 2012, respectively. He is currently working towards the Ph.D. degree at the same department. His current research interests include signal processing, wireless localization, and visible light communications.



Osman Erdem received the B.S. and M.S. degrees from the Department of Electrical and Electronics Engineering at Bilkent University, Ankara, Turkey, in 2015 and 2018, respectively. He is currently pursuing the Ph.D. degree in the same department. His current research interests include the fields of visible light communication and visible light positioning.



Sinan Gezici (S'03–M'06–SM'11) received the B.S. degree from Bilkent University, Turkey in 2001, and the Ph.D. degree in Electrical Engineering from Princeton University in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA. Since 2007, he has been with the Department of Electrical and Electronics Engineering at Bilkent University, where he is currently a Professor. Dr. Gezici's research interests are in the areas of detection and estimation theory, wireless communications, localization systems, and visible light communications.

Among his publications in these areas is the book *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols* (Cambridge University Press, 2008). Dr. Gezici has been an associate editor for *IEEE Transactions on Communications* and *IEEE Wireless Communications Letters*.

SUPPLEMENTARY MATERIAL FOR
 COOPERATIVE LOCALIZATION IN HYBRID
 INFRARED/VISIBLE LIGHT NETWORKS: THEORETICAL
 LIMITS AND DISTRIBUTED ALGORITHMS
 MUSA FURKAN KESKIN, OSMAN ERDEM AND SINAN GEZICI

S-I. PROOF OF LEMMA 1

Suppose that $\mathbf{x}_1 \in \mathcal{B}_\alpha$, $\mathbf{x}_2 \in \mathcal{B}_\alpha$, and $\alpha < \gamma$. It is clear that for any $\lambda \in (0, 1)$, $\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in \Omega$. Also, for any $\lambda \in (0, 1)$,

$$g_\epsilon(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \quad (\text{S1})$$

$$= \gamma - \frac{\lambda(\mathbf{y} - \mathbf{x}_1)^T \mathbf{n}_R + (1 - \lambda)(\mathbf{y} - \mathbf{x}_2)^T \mathbf{n}_R}{\|\lambda(\mathbf{x}_1 - \mathbf{y}) + (1 - \lambda)(\mathbf{x}_2 - \mathbf{y})\|^k + \epsilon} \quad (\text{S2})$$

$$\leq \gamma - \frac{\lambda(\mathbf{y} - \mathbf{x}_1)^T \mathbf{n}_R + (1 - \lambda)(\mathbf{y} - \mathbf{x}_2)^T \mathbf{n}_R}{\lambda \|\mathbf{x}_1 - \mathbf{y}\|^k + (1 - \lambda) \|\mathbf{x}_2 - \mathbf{y}\|^k + \epsilon} \quad (\text{S3})$$

$$\leq \gamma -$$

$$\frac{\lambda(\gamma - \alpha)(\|\mathbf{x}_1 - \mathbf{y}\|^k + \epsilon) + (1 - \lambda)(\gamma - \alpha)(\|\mathbf{x}_2 - \mathbf{y}\|^k + \epsilon)}{\lambda(\|\mathbf{x}_1 - \mathbf{y}\|^k + \epsilon) + (1 - \lambda)(\|\mathbf{x}_2 - \mathbf{y}\|^k + \epsilon)} \quad (\text{S4})$$

$$= \alpha \quad (\text{S5})$$

is obtained, where (S3) is due to the convexity of $\|\cdot\|^k$, $\mathbf{x}_1 \in \Omega$, and $\mathbf{x}_2 \in \Omega$, and (S4) follows from $\mathbf{x}_1 \in \mathcal{B}_\alpha$ and $\mathbf{x}_2 \in \mathcal{B}_\alpha$. Hence, (S1)–(S5) implies the convexity of \mathcal{B}_α for $\alpha < \gamma$. For $\alpha \geq \gamma$, $g_\epsilon(\mathbf{x}) \leq \alpha$ is satisfied $\forall \mathbf{x} \in \Omega$, which implies $\mathcal{B}_\alpha = \Omega$. Therefore, \mathcal{B}_α is convex $\forall \alpha \in \mathbb{R}$.

S-II. PROOF OF PROPOSITION 1

Since $\Lambda_j \cap \Upsilon_j \neq \emptyset$, consider any point $\mathbf{x}_j \in \Lambda_j \cap \Upsilon_j$. At the n th iteration, it can be assumed that $\mathbf{x}_j^{(n)} \notin \Lambda_j \cap \Upsilon_j$ because otherwise iterations will stop via (54) and (38), which implies quasi-Fejér convergence of $\{\mathbf{x}_j^{(n)}\}_{n=0}^\infty$ to $\Lambda_j \cap \Upsilon_j$ based on Definition 3. Then, based on the iterative step in (54), the following is obtained:

$$\|\mathbf{x}_j^{(n+1)} - \mathbf{x}_j\|^2 = \|\tilde{\mathbf{x}}_j^{(n)} - \mathbf{x}_j - \lambda_j^{(n)} \boldsymbol{\theta}_j^{(n)}\|^2 \quad (\text{S6})$$

where

$$\boldsymbol{\theta}_j^{(n)} \triangleq \sum_{k=1}^{K_j} \left(\sum_{\ell \in \tilde{S}_k^{(j)}} \tilde{\kappa}_{\ell,k}^{(j)} H_{\tilde{g}_{\ell,k}^{(j)}}(\tilde{\mathbf{x}}_j^{(n)}) + \sum_{i=1, i \neq j}^{N_V} \sum_{\ell \in S_k^{(i,j)}} \kappa_{\ell,k}^{(i,j)} H_{g_{\ell,k}^{(i,j)}(\cdot, \mathbf{x}_i^{(n)})}(\tilde{\mathbf{x}}_j^{(n)}) \right) \quad (\text{S7})$$

with the scaled gradient operator being defined as

$$H_f(\mathbf{x}) = \frac{f^+(\mathbf{x})}{\|\nabla f(\mathbf{x})\|^2} \nabla f(\mathbf{x}). \quad (\text{S8})$$

From (S6), it follows that

$$\begin{aligned} \|\mathbf{x}_j^{(n+1)} - \mathbf{x}_j\|^2 &= \|\tilde{\mathbf{x}}_j^{(n)} - \mathbf{x}_j\|^2 + \left(\lambda_j^{(n)}\right)^2 \|\boldsymbol{\theta}_j^{(n)}\|^2 \\ &\quad - 2\lambda_j^{(n)} \left(\boldsymbol{\theta}_j^{(n)}\right)^T \left(\tilde{\mathbf{x}}_j^{(n)} - \mathbf{x}_j\right). \end{aligned} \quad (\text{S9})$$

Let $\tilde{\mathcal{F}}_j$ and \mathcal{F}_j be the sets of Lambertian functions for the j th VLC unit corresponding to the noncooperative and cooperative cases, respectively, which are given by

$$\tilde{\mathcal{F}}_j = \left\{ \left\{ \tilde{g}_{\ell,k}^{(j)} \right\}_{\ell \in \tilde{S}_k^{(j)}} \right\}_{k \in \{1, 2, \dots, K_j\}} \quad (\text{S10})$$

$$\mathcal{F}_j = \left\{ \left\{ \left\{ g_{\ell,k}^{(i,j)}(\cdot, \mathbf{x}_i^{(n)}) \right\}_{\ell \in S_k^{(i,j)}} \right\}_{i \in \{1, 2, \dots, N_V\} \setminus j} \right\}_{k \in \{1, 2, \dots, K_j\}} \quad (\text{S11})$$

For any function $f \in \tilde{\mathcal{F}}_j \cup \mathcal{F}_j$, $f(\mathbf{x}_j) \leq 0$ holds since $\mathbf{x}_j \in \Lambda_j \cap \Upsilon_j$ (see (20), (22), (25), and (26)). Consider the following mutually exclusive and exhaustive subsets of $\tilde{\mathcal{F}}_j \cup \mathcal{F}_j$:

$$\mathcal{F}_{j,n}^* = \{f \in \tilde{\mathcal{F}}_j \cup \mathcal{F}_j \mid f(\tilde{\mathbf{x}}_j^{(n)}) \leq 0\} \quad (\text{S12})$$

$$\mathcal{F}_{j,n}^\circ = \{f \in \tilde{\mathcal{F}}_j \cup \mathcal{F}_j \mid f(\tilde{\mathbf{x}}_j^{(n)}) > 0\}. \quad (\text{S13})$$

It is clear from (S8) that for any $f^* \in \mathcal{F}_{j,n}^*$,

$$H_{f^*}(\tilde{\mathbf{x}}_j^{(n)}) = 0 \quad (\text{S14})$$

is satisfied. On the other hand, for any $f^\circ \in \mathcal{F}_{j,n}^\circ \cap \tilde{\mathcal{F}}_j$, $f^\circ(\mathbf{x}_j) = 0 < f^\circ(\tilde{\mathbf{x}}_j^{(n)})$, and, for any $f^\circ \in \mathcal{F}_{j,n}^\circ \cap \mathcal{F}_j$, $f^\circ(\mathbf{x}_j) < f^\circ(\tilde{\mathbf{x}}_j^{(n)})$ via Assumption A1. Then, the following inequality holds for any $f^\circ \in \mathcal{F}_{j,n}^\circ$:

$$f^\circ(\mathbf{x}_j) < f^\circ(\tilde{\mathbf{x}}_j^{(n)}). \quad (\text{S15})$$

Since \mathbf{x}_j and $\tilde{\mathbf{x}}_j^{(n)}$ both lie inside the halfspaces of the form (41) and (49) corresponding to the set of functions $\mathcal{F}_{j,n}^\circ$ ($\mathbf{x}_j \in \Lambda_j \subset \Gamma_j$ and $\tilde{\mathbf{x}}_j^{(n)} \in \Gamma_j$, see (40), (41), (53) and (25)), they are in the region where any $f^\circ \in \mathcal{F}_{j,n}^\circ$ is quasiconvex. Hence, from (S15) and Definition 2,

$$(\nabla f^\circ(\tilde{\mathbf{x}}_j^{(n)}))^T (\mathbf{x}_j - \tilde{\mathbf{x}}_j^{(n)}) \leq 0 \quad (\text{S16})$$

follows, which, based on (S8), implies that

$$\left(H_{f^\circ}(\tilde{\mathbf{x}}_j^{(n)})\right)^T \left(\tilde{\mathbf{x}}_j^{(n)} - \mathbf{x}_j\right) \geq 0. \quad (\text{S17})$$

The inner product term in (S9) can be decomposed into two parts corresponding to the sets $\mathcal{F}_{j,n}^*$ and $\mathcal{F}_{j,n}^\circ$. The part that corresponds to $\mathcal{F}_{j,n}^*$ is 0 via (S14) and the remaining part is greater than or equal to 0 via (S17). Hence, the following inequality is obtained:

$$\left(\boldsymbol{\theta}_j^{(n)}\right)^T \left(\tilde{\mathbf{x}}_j^{(n)} - \mathbf{x}_j\right) \geq 0, \quad (\text{S18})$$

which, based on (S9), yields

$$\|\mathbf{x}_j^{(n+1)} - \mathbf{x}_j\|^2 \leq \|\tilde{\mathbf{x}}_j^{(n)} - \mathbf{x}_j\|^2 + \epsilon_j^{(n)} \quad (\text{S19})$$

where

$$\epsilon_j^{(n)} \triangleq \left(\lambda_j^{(n)}\right)^2 \|\boldsymbol{\theta}_j^{(n)}\|^2. \quad (\text{S20})$$

From the fact that $\mathbf{x}_j \in \Gamma_j$, the following can be written:

$$\|\tilde{\mathbf{x}}_j^{(n)} - \mathbf{x}_j\| = \|P_{\Gamma_j}(\mathbf{x}_j^{(n)}) - P_{\Gamma_j}(\mathbf{x}_j)\| \quad (\text{S21})$$

$$\leq \|\mathbf{x}_j^{(n)} - \mathbf{x}_j\| \quad (\text{S22})$$

where (S21) follows from (53), and (S22) is due to the non-expansivity of the orthogonal projection operator. Combining (S21) and (S22) with (S19) yields the following inequality:

$$\|\mathbf{x}_j^{(n+1)} - \mathbf{x}_j\|^2 \leq \|\mathbf{x}_j^{(n)} - \mathbf{x}_j\|^2 + \epsilon_j^{(n)}. \quad (\text{S23})$$

Based on the parallel projection step in (54), it can easily be shown that

$$\|\mathbf{x}_j^{(n+1)} - \tilde{\mathbf{x}}_j^{(n)}\| = \lambda_j^{(n)} \|\boldsymbol{\theta}_j^{(n)}\| \quad (\text{S24})$$

where $\boldsymbol{\theta}_j^{(n)}$ is given by (S7). Then, from Assumption A2, it follows that $\sum_{n=0}^\infty \|\mathbf{x}_j^{(n+1)} - \tilde{\mathbf{x}}_j^{(n)}\|^2 < \infty$, which leads to $\sum_{n=0}^\infty \epsilon_j^{(n)} < \infty$ via (S24) and (S20). Finally, using (S23) and Definition 3 yields the desired result.

S-III. PROOF OF PROPOSITION 2

Following the same steps as stated in the proof of Proposition 1, the following inequality is obtained based on (50):

$$\|\mathbf{x}_j^{(n+1)} - \mathbf{x}_j\|^2 \leq \|\mathbf{x}_j^{(n)} - \mathbf{x}_j\|^2 + \epsilon_j^{(n)} \quad (\text{S25})$$

where

$$\epsilon_j^{(n)} \triangleq \left\| \vartheta_{\text{nc}} \lambda_{j,\text{nc}}^{(n)} H_{g_{\tilde{\ell}_{\text{nc}}, \tilde{k}_{\text{nc}}}}(\tilde{\mathbf{x}}_j^{(n)}) + \vartheta_c \lambda_{j,c}^{(n)} H_{g_{\tilde{\ell}_c, \tilde{k}_c}(\cdot, \mathbf{x}_i^{(n)})}}(\tilde{\mathbf{x}}_j^{(n)}) \right\|^2 \quad (\text{S26})$$

with H_f being defined as in (S8). The averaging step in (50) leads to $\|\mathbf{x}_j^{(n+1)} - \tilde{\mathbf{x}}_j^{(n)}\| = \sqrt{\epsilon_j^{(n)}}$, where $\epsilon_j^{(n)}$ is given by (S26). Assuming that A2 holds and following an approach similar to that in the proof of Proposition 1, the inequality $\sum_{n=0}^{\infty} \epsilon_j^{(n)} < \infty$ is obtained, thus establishing the quasi-Fejér convergence of the sequence $\{\mathbf{x}_j^{(n)}\}_{n=0}^{\infty}$ to $\Lambda_j \cap \Upsilon_j$.

S-IV. PROOF OF LEMMA 3

The proof is based on contradiction. Suppose that $\lim_{n \rightarrow \infty} \lambda_j^{(n)} = 0$. Then, for each $\zeta > 0$, there exists an iteration index $n(\zeta)$ such that $\lambda_j^{(n(\zeta))} < \zeta$. Based on the Armijo step size selection rule (44) in Algorithm 2 and the step size update equation (56) in Algorithm 4, there exists a function $f^\circ \in \tilde{\mathcal{F}}_j \cup \mathcal{F}_j$ such that the inequality in (44) is not satisfied for the step size

$$\tilde{\zeta} = \lambda_j^{(n(\zeta)-1)} \xi^{\bar{m}} \quad (\text{S27})$$

where

$$\bar{m} = \max\{m \in \mathbb{Z}_{\geq 0} \mid \lambda_j^{(n(\zeta)-1)} \xi^m > \zeta\}. \quad (\text{S28})$$

Hence, the following inequality is obtained:

$$f^\circ(G_{f^\circ}^{\tilde{\zeta}}(\tilde{\mathbf{x}}_j^{(n(\zeta))})) > f^\circ(\tilde{\mathbf{x}}_j^{(n(\zeta))})(1 - \beta\tilde{\zeta}). \quad (\text{S29})$$

It is clear that $f^\circ(\tilde{\mathbf{x}}_j^{(n(\zeta))}) > 0$ since otherwise the step size selection procedure would not need to be applied, meaning that $\tilde{\mathbf{x}}_j^{(n(\zeta))}$ is inside the zero-sublevel set of every function in $\tilde{\mathcal{F}}_j \cup \mathcal{F}_j$, i.e., $\tilde{\mathbf{x}}_j^{(n(\zeta))} \in \Lambda_j \cap \Upsilon_j$, which completes the proof of convergence of the iterates $\{\mathbf{x}_j^{(n)}\}_{n=0}^{\infty}$ to the set Λ_j via Lemma 2. Then, the left hand side of (S29) can be rewritten using the Taylor series expansion and (39) as follows:

$$f^\circ(G_{f^\circ}^{\tilde{\zeta}}(\tilde{\mathbf{x}}_j^{(n(\zeta))})) = f^\circ(\tilde{\mathbf{x}}_j^{(n(\zeta))}) - \tilde{\zeta} f^\circ(\tilde{\mathbf{x}}_j^{(n(\zeta))}) + O(\tilde{\zeta}^2) \quad (\text{S30})$$

where $O(\tilde{\zeta}^2)$ represents the terms with $\tilde{\zeta}^s$ for $s \geq 2$. Since $\lim_{\tilde{\zeta} \rightarrow 0} O(\tilde{\zeta}^2)/\tilde{\zeta} = 0$, there exists $v > 0$ such that

$$O(\tilde{\zeta}^2)/\tilde{\zeta} < f^\circ(\tilde{\mathbf{x}}_j^{(n(\zeta))})(1 - \beta) \quad (\text{S31})$$

is satisfied for $0 < \tilde{\zeta} \leq v$. The existence of v satisfying (S31) is guaranteed by $f^\circ(\tilde{\mathbf{x}}_j^{(n(\zeta))}) > 0$ and $\beta \in (0, 1)$. Inserting (S31) into (S30) yields the inequality

$$f^\circ(G_{f^\circ}^{\tilde{\zeta}}(\tilde{\mathbf{x}}_j^{(n(\zeta))})) < f^\circ(\tilde{\mathbf{x}}_j^{(n(\zeta))})(1 - \beta\tilde{\zeta}), \quad (\text{S32})$$

which contradicts with (S29). Therefore, the initial assumption is not valid, which implies $\lim_{n \rightarrow \infty} \lambda_j^{(n)} > 0$.

S-V. PROOF OF LEMMA 4

The proof can be obtained by invoking similar arguments to those in the proof of Lemma 3 and using the step size update rule in (51) and (52).

S-VI. PROOF OF PROPOSITION 3

From Proposition 1, $\sum_{n=0}^{\infty} \epsilon_j^{(n)} < \infty$, where $\epsilon_j^{(n)}$ is given by (S20) in the supplementary material. Hence, $\lim_{n \rightarrow \infty} \epsilon_j^{(n)} = 0$ is obtained. Based on Lemma 3, (S7), and (S20), it follows that

$$\lim_{n \rightarrow \infty} \left\| \sum_{k=1}^{K_j} \left(\sum_{\ell \in \tilde{S}_k^{(j)}} \tilde{\kappa}_{\ell,k}^{(j)} H_{g_{\ell,k}}(\tilde{\mathbf{x}}_j^{(n)}) + \sum_{i=1, i \neq j}^{N_V} \sum_{\ell \in S_k^{(i,j)}} \kappa_{\ell,k}^{(i,j)} H_{g_{\ell,k}}(\cdot, \mathbf{x}_i^{(n)}) (\tilde{\mathbf{x}}_j^{(n)}) \right) \right\| = 0, \quad (\text{S33})$$

which implies that

$$\lim_{n \rightarrow \infty} J_f(\tilde{\mathbf{x}}_j^{(n)}) = 0 \quad (\text{S34})$$

is satisfied $\forall f \in \tilde{\mathcal{F}}_j \cup \mathcal{F}_j$, where the operator J_f defined on \mathbb{R}^d for the set of continuously differentiable functions $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is given by

$$J_f(\mathbf{x}) = \left(\frac{f^+(\mathbf{x})}{\|\nabla f(\mathbf{x})\|} \right)^2. \quad (\text{S35})$$

For a generic Lambertian function in (28), the norm square of the gradient can be expressed as

$$\|\nabla g_\epsilon(\mathbf{x})\|^2 = \frac{1}{(\|\mathbf{x} - \mathbf{y}\|^k + \epsilon)^2} + \left(\frac{(\mathbf{y} - \mathbf{x})^T \mathbf{n}_R}{\|\mathbf{x} - \mathbf{y}\|^k + \epsilon} \right)^2 \frac{k \|\mathbf{x} - \mathbf{y}\|^{k-2} ((k-2)\|\mathbf{x} - \mathbf{y}\|^k - 2\epsilon)}{(\|\mathbf{x} - \mathbf{y}\|^k + \epsilon)^2}. \quad (\text{S36})$$

Since the sequence of iterates $\{\tilde{\mathbf{x}}_j^{(n)}\}_{n=0}^{\infty}$ is bounded by Lemma 2, $\{\|\tilde{\mathbf{x}}_j^{(n)} - \mathbf{y}\|\}_{n=0}^{\infty}$ is also bounded, which implies the boundedness of $\|\nabla g_\epsilon(\mathbf{x})\|$ via (S36). Therefore, based on (S34) and (S35), it follows that

$$\lim_{n \rightarrow \infty} f^+(\tilde{\mathbf{x}}_j^{(n)}) = 0, \quad \forall f \in \tilde{\mathcal{F}}_j \cup \mathcal{F}_j. \quad (\text{S37})$$

From the Bolzano-Weierstrass Theorem [1, Section 3.4], the boundedness of $\{\tilde{\mathbf{x}}_j^{(n)}\}_{n=0}^{\infty}$ requires that $\{\tilde{\mathbf{x}}_j^{(n)}\}_{n=0}^{\infty}$ has a convergent subsequence. Denote the limit of this subsequence by \mathbf{x}_j^* . From (S37), it turns out that $\mathbf{x}_j^* \in \Lambda_j \cap \Upsilon_j$. Therefore, $\Lambda_j \cap \Upsilon_j$ contains a limit point of $\{\tilde{\mathbf{x}}_j^{(n)}\}_{n=0}^{\infty}$, which, based on Lemma 2, yields the result that $\{\tilde{\mathbf{x}}_j^{(n)}\}_{n=0}^{\infty}$ converges to a point inside $\Lambda_j \cap \Upsilon_j$. Based on (53) and the fact that $\Lambda_j \subset \Gamma_j$ (see (25) and (40)), it follows that the sequence $\{\mathbf{x}_j^{(n)}\}_{n=0}^{\infty}$ converges to a point $\mathbf{x}_j \in \Lambda_j \cap \Upsilon_j$.

S-VII. PROOF OF PROPOSITION 4

Applying similar steps to those in the proof of Proposition 3 and exploiting Proposition 2 and Lemma 4, the following results are obtained:

$$\lim_{n \rightarrow \infty} \left[\tilde{g}_{\hat{\ell}_{nc}, \hat{k}_{nc}}^{(j)}(\tilde{\mathbf{x}}_j^{(n)}) \right]^+ = 0 \quad (\text{S38})$$

$$\lim_{n \rightarrow \infty} \left[g_{\hat{\ell}_c, \hat{k}_c}^{(i_c, j)}(\tilde{\mathbf{x}}_j^{(n)}, \mathbf{x}_i^{(\hat{n})}) \right]^+ = 0 \quad (\text{S39})$$

Based on the most violated constraint control in (46) and (47), it is obvious that

$$\tilde{f}(\tilde{\mathbf{x}}_j^{(n)}) \leq \tilde{g}_{\hat{\ell}_{nc}, \hat{k}_{nc}}^{(j)}(\tilde{\mathbf{x}}_j^{(n)}), \quad \forall \tilde{f} \in \tilde{\mathcal{F}}_j \quad (\text{S40})$$

$$f(\tilde{\mathbf{x}}_j^{(n)}) \leq g_{\hat{\ell}_c, \hat{k}_c}^{(i_c, j)}(\tilde{\mathbf{x}}_j^{(n)}, \mathbf{x}_i^{(\hat{n})}), \quad \forall f \in \mathcal{F}_j \quad (\text{S41})$$

which implies via (S38) and (S39) that

$$\lim_{n \rightarrow \infty} f^+(\tilde{\mathbf{x}}_j^{(n)}) = 0, \quad \forall f \in \tilde{\mathcal{F}}_j \cup \mathcal{F}_j. \quad (\text{S42})$$

The rest of the proof is the same as that in Proposition 3.

S-VIII. PARTIAL DERIVATIVES IN (12)

From (5) and (6), the partial derivatives in (12) are obtained as follows:

$$\begin{aligned} \frac{\partial \tilde{\alpha}_{l,k}^{(j)}(\mathbf{x}_j)}{\partial x_t} &= - \frac{(\tilde{m}_l + 1) \tilde{P}_{T,l} A_k^{(j)} ((\tilde{\mathbf{d}}_{l,k}^{(j)})^T \tilde{\mathbf{n}}_{T,l})^{\tilde{m}_l}}{2\pi \|\tilde{\mathbf{d}}_{l,k}^{(j)}\|^{\tilde{m}_l+3}} \\ &\times \left(\tilde{m}_l \tilde{n}_{T,l}(t-3j+3) (\tilde{\mathbf{d}}_{l,k}^{(j)})^T \mathbf{n}_{R,k}^{(j)} ((\tilde{\mathbf{d}}_{l,k}^{(j)})^T \tilde{\mathbf{n}}_{T,l})^{-1} \right. \\ &\quad \left. + n_{R,k}^{(j)}(t-3j+3) \right) \\ &\quad - (\tilde{m}_l + 3) \tilde{d}_{l,k}^{(j)}(t-3j+3) (\tilde{\mathbf{d}}_{l,k}^{(j)})^T \mathbf{n}_{R,k}^{(j)} \|\tilde{\mathbf{d}}_{l,k}^{(j)}\|^{-2} \end{aligned} \quad (\text{S43})$$

for $t \in \{3j-2, 3j-1, 3j\}$ and $\partial \tilde{\alpha}_{l,k}^{(j)}(\mathbf{x}_j)/\partial x_t = 0$ otherwise, where $\tilde{n}_{T,l}(t-3j+3)$, $n_{R,k}^{(j)}(t-3j+3)$, and $\tilde{d}_{l,k}^{(j)}(t-3j+3)$ represent the $(t-3j+3)$ th elements of $\tilde{\mathbf{n}}_{T,l}$, $\mathbf{n}_{R,k}^{(j)}$, and $\tilde{\mathbf{d}}_{l,k}^{(j)}$, respectively. Similarly,

$$\begin{aligned} \frac{\partial \alpha_{l,k}^{(i,j)}(\mathbf{x}_j, \mathbf{x}_i)}{\partial x_t} &= - \frac{(m_l^{(i)} + 1) P_{T,l}^{(i)} A_k^{(j)} ((\mathbf{d}_{l,k}^{(i,j)})^T \mathbf{n}_{T,l}^{(i)})^{m_l^{(i)}}}{2\pi \|\mathbf{d}_{l,k}^{(i,j)}\|^{m_l^{(i)}+3}} \\ &\times \left(m_l^{(i)} n_{T,l}^{(i)}(t-3j+3) (\mathbf{d}_{l,k}^{(i,j)})^T \mathbf{n}_{R,k}^{(j)} ((\mathbf{d}_{l,k}^{(i,j)})^T \mathbf{n}_{T,l}^{(i)})^{-1} \right. \\ &\quad \left. + n_{R,k}^{(j)}(t-3j+3) \right) \\ &\quad - (m_l^{(i)} + 3) d_{l,k}^{(i,j)}(t-3j+3) (\mathbf{d}_{l,k}^{(i,j)})^T \mathbf{n}_{R,k}^{(j)} \|\mathbf{d}_{l,k}^{(i,j)}\|^{-2} \end{aligned} \quad (\text{S44})$$

for $t \in \{3j-2, 3j-1, 3j\}$, $\partial \alpha_{l,k}^{(i,j)}(\mathbf{x}_i, \mathbf{x}_j)/\partial x_t$ is equal to the negative of (S44) with $(t-3j+3)$'s being replaced by $(t-3i+3)$'s for $t \in \{3i-2, 3i-1, 3i\}$, and $\partial \alpha_{l,k}^{(i,j)}(\mathbf{x}_i, \mathbf{x}_j)/\partial x_t = 0$ otherwise. In (S44), $n_{T,l}^{(i)}(t-3j+3)$ and $d_{l,k}^{(i,j)}(t-3j+3)$ denote the $(t-3j+3)$ th elements of $\mathbf{n}_{T,l}^{(i)}$ and $\mathbf{d}_{l,k}^{(i,j)}$, respectively.

REFERENCES

- [1] R. G. Bartle and D. R. Sherbert, *Introduction to Real Analysis*. John Wiley & Sons, 2000.