Odometry-Aided mmWave Communications using Overparameterized Beamforming Optimization

Gary C.F. Lee
Communications and Networks
Institute for Infocomm Research
Singapore, Singapore
gary_lee@i2r.a-star.edu.sg

Ernest Kurniawan
Communications and Networks
Institute for Infocomm Research
Singapore, Singapore
ekurniawan@i2r.a-star.edu.sg

Yuanjin Zheng
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore, Singapore
yjzheng@ntu.edu.sg

***Abstract*—In vehicular-to-infrastructure (V2I) applications, as we adopt higher frequencies and antenna array processing for faster data rates, beamforming is a critical enabling technology to preserve high-quality communication links between mobile users and the base station. Traditional beamforming requires overhead in establishing the directional link, compounded by short coherence time of mobile users, which necessitates frequent re-calibration. Recent works consider more efficient strategies by leveraging context information, with position-aware approaches showing particular promise. Nonetheless, uncertainties in even the most advanced localization and odometry methods can lead to compounded errors in position estimation for downstream tasks such as beam alignment and tracking. Our work addresses these challenges by factoring in the uncertainties inherent in practical odometry systems. We propose a beamforming optimization that accounts for this uncertainty, complemented by overparameterization strategies in gradient-based optimization to the non-linear problem of interest. Our simulation results demonstrate the feasibility of maintaining robust communication links in representative scenarios by leveraging localization and odometry information, despite the challenges posed by a mobile platform with imprecise position estimates.**

***Index Terms*—beamforming, odometry, beam tracking, beam pattern optimization, mmWave communications.**

I. Introduction

As the technology for autonomous vehicles and assisted driving advances, enabled by cutting-edge sensors and wireless technology, our research work is motivated by ensuring seamless vehicular-to-infrastructure (V2I) connectivity. Central to this is millimeter-wave (mmWave) communications, relevant to the latest standards in 5G and WiFi, and crucial for achieving high data rates and low latency links [1], [2].

Beamforming is an important component for such modern wireless communication systems [3], [4]. It involves using an antenna array to direct energy towards particular directions, enhancing spatial selectivity, and counteracting the high free space path loss associated with higher carrier frequencies, such as those in the 5 GHz and millimeter-wave bands.

Conventional beamforming often necessitates calibration to establish knowledge about the channel state between the user equipment (UE) and the base station (BS). Other types of context information, such as relative positioning between the UE and BS, can be exploited, especially when channel state information is not readily available [5].

In this work, we investigate beamforming strategies leveraging the information that autonomous vehicles already generate and process. Specifically, we suppose that odometry and localization can be attained through the use of sensors and computation. How can these be used as contextual information for more efficient beamforming and robust beam tracking? A key practical challenge we address is the inherently imperfect position estimates. How should the beamforming strategy account for the position inaccuracies to robustify the beamforming scheme? Our work posits that uncertain position estimates require wider beam patterns, which, although potentially resulting in slightly reduced path gains, offer broader and more robust coverage. We investigate this relationship, proposing a numerical optimization approach to address these challenges.

Notation: Vectors and matrices are denoted by boldface lowercase and uppercase letters, respectively (e.g. $\boldsymbol{g}$ and $\boldsymbol{G}$). $\boldsymbol{G}^{\mathrm{T}}$ and $\boldsymbol{G}^{\mathrm{H}}$ denotes the transpose and Hermitian of $\boldsymbol{G}$ respectively. $|g|$ denotes the absolute value of scalar g, or equivalently, its magnitude for complex-valued g.

II. Relevant Works

Beam training is a prevalent strategy that involves the transmitter exhaustively searching over a pre-defined set of beamforming vectors to identify the optimal one based on some criteria (e.g. maximizing signal-to-noise ratio (SNR)). However, this process can be slow due to the exponentially growing codebook size with the number of antenna elements and desired angular precision. Its effectiveness also depends on the channel remaining unchanged (quasi-static) for an extended period, which is not always the case with mobile user. Consequently, frequent beam training is required to adapt to changing channel conditions, which, in turn, affects the data rate. To this end, we look to other beamforming techniques, particularly by leveraging information about the UEs.

Adaptive beamforming is another technique aimed at enhancing the received signal quality [6], relying on access to statistics of the channel and received signal. However, extracting the relevant channel information also typically requires overheads in the form of coordination or control messages.

This research is supported by the National Research Foundation, Singapore and Infocomm Media Development Authority under its Future Communications Research & Development Programme (FCP-NTU-RG-2022-021). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of National Research Foundation, Singapore and Infocomm Media Development Authority.

On the other hand, recent studies have focused on using context information, such as position information, to improve the effectiveness of beamforming [5].

Of particular interest is the use of in-vehicle sensors, such as inertial measurement unit (IMU), stereo cameras and light detection and ranging (LiDAR) sensors, to aid in beamforming optimization. These sensors are already integral to many modern autonomous systems, eliminating the need for additional specialized equipment. Some works [7]–[9] have proposed neural network solutions that learn the correlation between raw sensor data (e.g. LiDAR scans) and the optimal beamforming weights. Such methods, however, require extensive labeled datasets for training, and might lack generalizability to unseen and unfamiliar environments. Our approach is agnostic to specific localization and odometry methods used by vehicular systems, focusing instead on utilizing position estimates and their associated uncertainties.

In relevant works on beamforming strategies and optimization, we recognize a collection of works related to codebook optimization. Conventional approaches, based on the assumption of uniform spatial distribution of devices, lead to codebooks that are divided into evenly-spaced angular sectors, utilizing a discrete Fourier transform (DFT)-based weights—in line with the maximum ratio combining strategy in line-of-sight (LOS) scenarios. Recognizing the sparsity of user distribution, [10] proposes a novel codebook design that combines narrow beam to create wider ones, using a novel objective function and a gradient-based optimization approach. Similarly, [11] proposes the use of complex-valued neural networks to learn a new codebook, although this requires channel-state information for training. Both methods focus on codebook design rather than tracking a moving source.

Robustness to measurement uncertainties can be attained with Bayesian beamforming [12], [13], achieves through a weighted sum of different spatial Wiener beamformers based on the posterior over the corresponding UE directions. However, the resulting beamforming weights may not conform to practical implementation constraints such as constant modulus requirements for phase-shifting-only implementations. The latter constraint induces a non-convex problem—one that is also tackled in aforementioned works [10], [11] with gradient-based approach, albeit for codebook optimization rather than beamforming to (and tracking) a single target.

III. Problem Formulation and Background

A. System Setup

Consider a single-antenna UE in the form of a mobile vehicle, which is communicating with a BS equipped with a K-element multi-antenna array. We focus on this multiple-input-single-output (MISO) link setting for this work[1].

The geometric channel model for the LOS scenario, the $K \times 1$ MISO channel vector between the BS and the UE (at the spherical coordinates (r, θ, φ)) can be described by

$$\boldsymbol{h} \triangleq \frac{\lambda_c}{4\pi r} \boldsymbol{a}(\theta, \varphi) \tag{1}$$

[1]or symmetrically, SIMO link when we consider the uplink scenario

where λ_c is the wavelength of the carrier wave, $r_{\text{UE-BS}}$ is the (LOS) distance between the UE and the BS, and $\boldsymbol{a}(\theta, \varphi)$ is the normalized steering vector towards the angles of departure (AOD), specifically the elevation and azimuth θ, φ. This complex-valued vector describes the change in phase and amplitude between the BS's antenna elements and UE.

The received signal at the UE can be expressed as

$$y = \sqrt{P} \boldsymbol{h}^{\mathrm{H}} \boldsymbol{w} s + z \tag{2}$$

where $s \in \mathbb{C}$ is the unit-power signal to be transmitted, P is the effective transmit power and $z \in \mathcal{CN}(0, \sigma^2_{\text{AWGN}})$ corresponds to the zero-mean additive white Gaussian noise (AWGN) term. The beamforming vector (also termed as precoding vector in this case) $\boldsymbol{w} \in \mathbb{C}^{K \times 1}$ corresponds to the phase shifts introduced at the BS to help steer the transmitted beam in the direction of the UE.

Considering the power constraints and practical considerations of the system, we restrict our attention to phase-only beamforming; in other words, weights are modeled as constant modulus, $|w_k| = 1/\sqrt{K}$ for all K elements. This setting is representative of antenna arrays realized by phase shifters.

This work is motivated by the directional nature of narrow beams at mmWave and with large antenna arrays, thereby the dependency on line-of-sight (LOS) paths. We focus on the LOS scenario, with some insights into non-LOS (NLOS) in our concluding remarks.

The objective is to find the optimal beamforming vector that maximizes receive power in the intended direction. This corresponds to finding $\boldsymbol{w}$ that maximizes the SNR at the UE, which we can express as

$$\boldsymbol{w}^{\text{opt}} = \arg\max_{\boldsymbol{w}} \frac{P|\boldsymbol{h}^{\mathrm{H}}\boldsymbol{w}|^2}{\sigma_{\text{AWGN}}} = \arg\max_{\boldsymbol{w}} |\boldsymbol{h}^{\mathrm{H}}\boldsymbol{w}|^2 \quad \text{s.t. } |w_k| = 1/\sqrt{K} \tag{3}$$

recalling $\boldsymbol{w}$ having a constant modulus constraint and being independent of P and σ_{AWGN}. Noting that $|\boldsymbol{h}^{\mathrm{H}}\boldsymbol{w}| \in \mathbb{C}$ is a scalar complex value, omitting the square still results in the same optimization solution.

Nevertheless, explicit knowledge of the UE's instantaneous channel state information (related to the desired AOD from the BS) can be difficult to obtain. Instead, we leverage existing odometry tools capable of estimating relative positions of UE from the BS, drawing its relation to the characteristics of the channel, and thereby result in obtaining optimal and robust beamforming weights $\boldsymbol{w}$.

B. Antenna Array Geometry

In this section, we further unpack the steering vector in the context of specific array geometries. We consider the antenna array to be either a uniform rectangular (planar) array (URA) or a uniform linear array (ULA)—for which we note that a ULA is a particular instance of URA with only one row of antenna elements. Let K_h and K_v be the number of horizontal and vertical elements, respectively, such that $K = K_h K_v$ (with $K_v = 1$ for ULA). URA has uniform spacing between antenna elements (both horizontally and vertically), which we take to be $d = \frac{\lambda_c}{2}$.

Suppose that the relative position of the UE, relative to the BS, can be represented by either the Cartesian coordinates (x, y, z) (the canonical representation in most localization algorithms) or the spherical coordinates (r, θ, φ) (the canonical representation for beamforming/radio-frequency sensing applications). Further, we suppose that the URA is oriented such that it is on the XZ-plane

$\boldsymbol{a}(\theta, \varphi)$, the array steering vector of such a K-element URA [14] as a function of the AOD of the above-mentioned UE, can be expressed as

$$\boldsymbol{a}(\theta,\varphi) = \frac{1}{\sqrt{K}}\left[1, \cdots, \exp\left(j\pi(k_v-1)\psi_v + j\pi(k_h-1)\psi_h\right), \cdots, \exp\left(j\pi(K_v-1)\psi_v + j\pi(K_h-1)\psi_h\right)\right]^{\mathrm{T}} \quad (4)$$

where ψ_v and ψ_v are related to the θ and φ by[2]

$$\psi_h = \sin\theta\cos\varphi = x/\sqrt{x^2+y^2+z^2}\ ;$$
$$\psi_v = \cos\theta = z/\sqrt{x^2+y^2+z^2},$$

where (x, y, z) are the Cartesian coordinates of the UE.

C. Position Information

We consider the case where additional information about the UE is available—specifically, that the UE is capable of performing a position estimate relative to the BS, for example, through IMU measurements or visual odometry (VO) with camera and LiDAR sensors. However, we also recognize inherent errors in this position estimate. Without loss of generality, we consider the BS as our frame's origin, and we model

$$x^*[n] = \widehat{x}[n]+\varepsilon_x\ ;\ y^*[n] = \widehat{y}[n]+\varepsilon_y\ ;\ z^*[n] = \widehat{z}[n]+\varepsilon_z \quad (5)$$

where $(x^*[n], y^*[n], z^*[n])$ is the true UE position at timestep n, and $(\widehat{x}[n], \widehat{y}[n], \widehat{z}[n])$ is the position estimate obtained by the relevant localization/odometry algorithm on the UE. As a result, there is some error, expressed in the Cartesian coordinates $(\varepsilon_x, \varepsilon_y, \varepsilon_z)$, whose characteristics would depend on the algorithm for the position estimate. We suppose that the position uncertainties may be represented in various ways. Broadly, we model the errors $(\varepsilon_x, \varepsilon_y, \varepsilon_z)$ as random variables, which follow some probability distribution. For instance, we consider two representative forms, namely,

1) a Gaussian distribution over the residual, described by zero-mean and covariance over the Cartesian axes; or
2) a region of confidence—this corresponds to a uniform distribution within the support described by the corresponding feasible region (e.g. a cuboid).

In the subsequent section, we present relevant background on relating relative positioning of BS and UE to beamforming optimization in the LOS scenario, and develop solutions that account for positioning uncertainty distributions.

[2]We retain the arguments of steering vector as θ and φ for ease of this exposition, recalling the dependency of ψ_v and ψ_v on these quantities.

IV. Methodology

A. Directional Beamforming

We first look at beamforming strategies once the position is known (e.g. through an oracle). In particular, the maximum SNR, or maximum ratio combining (MRC) solution for a single UE at (r, θ, φ), subject to constraint in (3), corresponds to

$$\boldsymbol{w}^{(\mathrm{MRC})} = \arg\max_{\boldsymbol{w}} \frac{\lambda_c}{4\pi r}|\boldsymbol{a}^{\mathrm{H}}(\theta,\varphi)\boldsymbol{w}| = \boldsymbol{a}(\theta,\varphi). \quad (6)$$

In the line-of-sight scenario, the solution is the vector in the Fourier dictionary corresponding to the LOS angle (azimuth and elevation) between the BS and the UE—i.e., (4).

B. Beamforming Optimization

In situations where precise information on r, θ, φ is not available, we adopt a probabilistic approach by modeling these parameters as random variables, as described in Subsection III-C. These errors are commonly expressed in Cartesian coordinates, or which we can perform the relevant transformations to spherical coordinates.

The proposed beamforming optimization formulation can thus be expressed as

$$\boldsymbol{w}^{(\mathrm{opt})} = \arg\max_{\boldsymbol{w}} \mathbb{E}_{r,\theta,\varphi}\left[\frac{\lambda_c}{4\pi r}|\boldsymbol{a}^{\mathrm{H}}(\theta,\varphi)\boldsymbol{w}|\right],. \quad (7)$$

Recalling the constant modulus constraint (3), this makes (7) a non-convex optimization problem. Similar to [10] and [11], we approach this with a gradient-based optimization after establishing the objective function. An important aspect of our approach is the decision to compute the expectation in terms of magnitudes (effectively akin to an $L1$ norm). Such an approach promotes high array gains uniformly across the region of interest, rather than allowing a few disproportionately large values to dominate and influence the solution, mitigating the instability mentioned by [10].

Departing from the previous works, which emphasize optimization based on channel state information, our work adopts a distinct system-level perspective. Central to our approach is how we utilize position estimate (which we assume is derived from the vehicular system's localization capabilities). We not only consider the estimated position but also factor in its associated uncertainties. With this information, we can compute the AOD and, in turn, the appropriate steering vectors for the LOS scenarios. These elements form the essential foundation of our proposed algorithm, which we detail in the following subsection.

C. Optimization Pipeline

We propose a Monte Carlo approximation of (7) and adopt a gradient-based optimization procedure, leading to Algorithm IV-C. Notably, our integrated approach begins with localization and uncertainty quantification, followed by geometric transformation to spherical coordinates to obtain AOD uncertainties. Subsequently, using Monte Carlo sampling, we perform a numerical gradient-based optimization to seek a solution for (7).

Algorithm 1 Gradient-based Monte Carlo Optimization

[At a given time step]

Input: $x_{\text{est}}, y_{\text{est}}, z_{\text{est}}, P_{\varepsilon_x}, P_{\varepsilon_y}, P_{\varepsilon_z}$, "Training" dataset size M, Total iterations T, Total additional trainable layers L

Output: Beamforming weights, $\boldsymbol{w}$

for $m = 0, 1, \cdots M-1$ **do**

Draw sample from distribution $\varepsilon_{(x,y,z)} \sim P_{\varepsilon_{(x,y,z)}}$;

$(x_m, y_m, z_m) = (x_{\text{est}} + \varepsilon_x, y_{\text{est}} + \varepsilon_y, z_{\text{est}} + \varepsilon_z)$

Transform coordinates $(x_m, y_m, z_m) \to (r_m, \theta_m, \varphi_m)$

Compute channel vector $\boldsymbol{h}_m = \frac{1}{r_m}\boldsymbol{a}(\theta_m, \varphi_m)$

(noting $\frac{\lambda_c}{4\pi}$ is a constant factor independent of m)

end for

Initialize trainable weights

$\Theta^{(0)} = \{\boldsymbol{D}_1^{(0)}, \boldsymbol{D}_2^{(0)}, \cdots \boldsymbol{D}_L^{(0)}, \boldsymbol{d}^{(0)}\}$

for $t = 0, 1, \cdots T-1$ **do**

$\boldsymbol{w}^{(t)} = \boldsymbol{D}_1^{(t)} \boldsymbol{D}_2^{(t)} \cdots \boldsymbol{D}_L^{(t)} \boldsymbol{d}^{(t)}$

$w_k^{(t)} = w_k^{(t)} / \sqrt{K} |w_k^{(t)}|$, $\forall k$

$\mathcal{L} = -\frac{1}{M} \sum_m |\boldsymbol{h}_m^{\text{H}} \boldsymbol{w}^{(t)}|$ (negative of objective (7))

Compute gradient update based on loss $\mathcal{L}$

Update parameters $\Theta^{(t)} \to \Theta^{(t+1)}$

end for

$\boldsymbol{w} = \boldsymbol{D}_1^{(T)} \boldsymbol{D}_2^{(T)} \cdots \boldsymbol{D}_L^{(T)} \boldsymbol{d}^{(T)}$

Expanding on the gradient-based optimization step, we remark that this process is akin to training a deep linear neural network, with the Monte Carlo samples serving as the "training set". This similarity suggests the potential benefits of adopting an overparameterization strategy, informed by recent advancements in deep learning techniques. That is, we can represent the vector $\boldsymbol{w} = \boldsymbol{D_1 D_2} \cdots \boldsymbol{D_L d}$, where the total number of trainable parameters is $LK^2 + K$. Rather counterintuitively, this increase in parameters, despite leading to increased computational load, actually accelerates the convergence rate towards an optimal solution—a phenomenon corroborated in recent works on overparameterization and convergence of neural networks [15].

To demonstrate this balance between computational time and convergence efficiency, we conduct an empirical characterization where we optimize the objective function for a 64-element ULA over a region of AODs. The additional fully connected layers lead to a notable speed up in convergence (in terms of number of gradient steps) to the optimal solution. However, the additional computational load of extra trainable parameters (i.e., time per gradient step) may negate the efficient convergence benefits. Fig. 1 charts the convergence against wallclock time under different values of L, where we observe $L = 3$ (introducing $3K^2$ additional parameters) offering a good balance between these two factors.

Equipped with the proposed method, we evaluate our beamforming optimization strategy in a few representative scenarios in the next section.

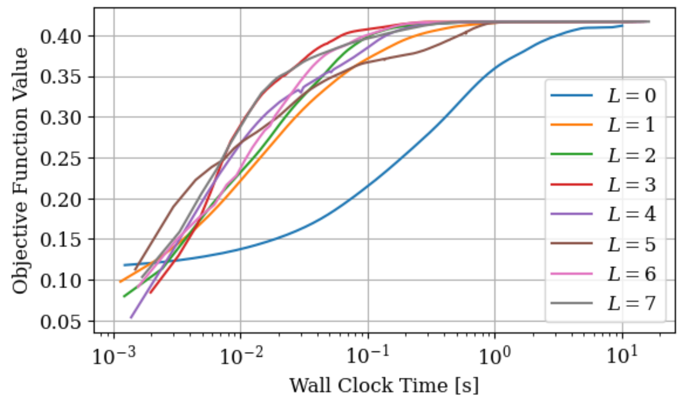


Fig. 1. Effects of overparameterization, plotting the iteration's objective function value against wallclock time. For this characterization, we perform gradient descent over $T = 10,000$ iterations. $L = 3$ leads to a good balance in depth, leading to the fastest convergence.

V. Experimental Details

A. Tools and Setup

For our beamforming optimization, we used PyTorch [16] to implement Algorithm IV-C, with $T = 1000$ and $L = 3$. We characterize our proposed approach by first characterizing the far-field beam pattern from our beamforming solutions; we then evaluate the overall performance by integrating these solutions with scene measurement data and odometry methods in representative simulations.

B. Beam Pattern Characterization

We begin by illustrating the far-field behavior of the beamforming weights by plotting the antenna array factor. This visualization reflects the directivity of the resulting beam pattern, depicting the expected signal magnitude as received from the respective directions. For this purpose, we consider a 64-element Uniform Linear Array (ULA).

Fig. 2 plots the beam pattern, characterized by the array factor given by the equation

$$A(\theta_0, \varphi_0) = |\boldsymbol{a}^{\text{H}}(\theta_0, \varphi_0)\boldsymbol{w}|. \tag{8}$$

The array factor describes the far-field array response using beamforming strategy $\boldsymbol{w}$, plotting the magnitude against direction. Note that while $\boldsymbol{w}$ is computed based on $\theta_{\text{est}}, \varphi_{\text{est}}$, we are interested in the gain at the actual UE's position $\theta_{\text{true}}, \varphi_{\text{true}}$, for which there may be a mismatch. For instance, if the true $\theta_{\text{true}} = 0°$, but the estimated $\theta_{\text{est}} = -3°$, the MRC beamformer would have a maximum magnitude at θ_{est}, but the effective magnitude would be significantly lower, around -13.3 dB at θ_{true}. In contrast, a robust beamformer accounts for the uncertainty in θ_{est}. Statistical characterization allows us to model the uncertainty, which can be introduced to the beamforming optimization, thereby obtaining beamforming weights that obtain a gain of around -6 dB instead.

The main advantage of such a optimization-based beamforming lies in its robustness to angle mismatch, which is crucial to prevent deterioration of array gain and potential loss of the communication link.

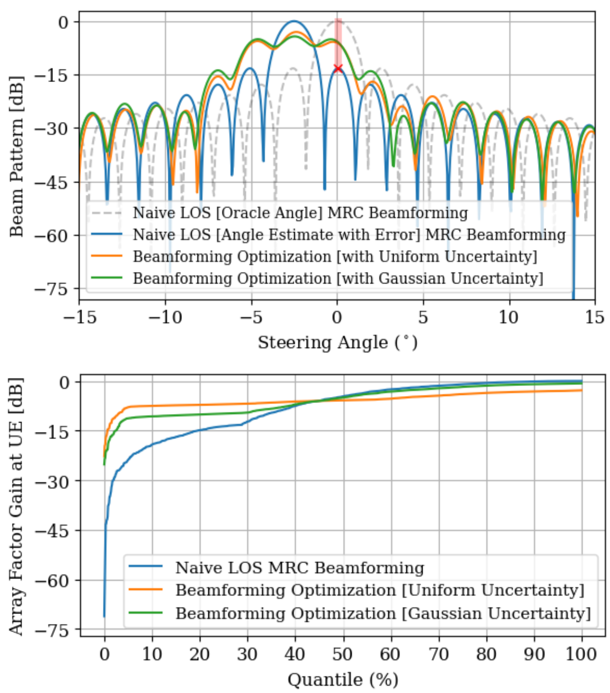


Fig. 2. Top: Array factor plot for different beamforming strategies based on the estimated position (corresponding to an AOD of around $-2.5°$), highlighting the decrease in gain at the true AOD (marked by the red bar); Bottom: Quantile plot of array factor gain at θ_{true}, illustrating the higher incidence of low gains with naive MRC compared to the more robust coverage achieved by our proposed beamforming method.

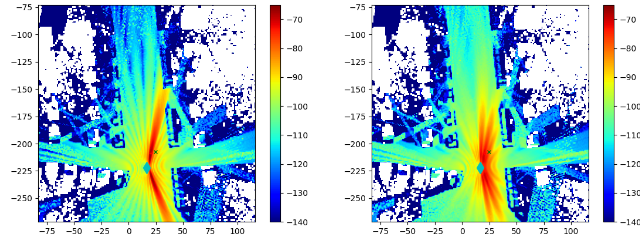
To demonstrate this, we conduct a Monte Carlo simulation with 10,000 realizations of noisy perturbations, assuming a Gaussian distribution with a standard deviation of 0.25 m along the x and y axes. In scenarios where the naive MRC beamformer is used based on perturbed positions, there is a significant probability of encountering low gains—i.e., gains fell below -15 dB in 20% of the cases. In contrast, with our proposed beamforming, this probability reduces to less than 3%. However, it should be noted that the maximum path gain achieved with our proposed method tends to be lower.

C. Demonstration with Odometry

To further evaluate our proposed method, we incorporate it into a system-level simulation, focusing on a realistic trajectory trace taken by a moving UE. We utilize odometry methods based on sensor measurements to estimate the UE's position, and assess performance by geometric channel modeling.

For this simulation, we consider a center frequency of 28 GHz. At this frequency, significant free space path loss occurs, which we counteract with an antenna array to electronically focus and steer the beam towards the UE. We consider both a 64-element ULA, and a $(16, 4)$-element URA for our simulation.

Using KITTI visual odometry dataset [17], we consider trajectory estimates obtained through two possible means—one, by integrating the northbound and eastbound velocities as measured through the IMU measurements; and two, by utilizing stereo camera and LiDAR recordings to perform VO, specifically the RGB-L SLAM algorithm [18], to obtain the real-time position estimate. In our simulation, we focus on the dataset's sequence 06, specifically between the 600th and 800th frames as the vehicle makes a U-turn. Based on statistical characterization of the algorithm, we model the position uncertainties as time-independent and uniformly distributed random variables with an interval of ± 4 m along the x-axis and ± 12 m along the y-axis, while the z-axis coordinate is assumed to be fixed at 1.5 m for simplicity.



(a) (b)

Fig. 3. Representative snapshot of a frame from the beamforming simulation using KITTI visual odometry dataset [17]. Coverage maps are generated with a (16,4)-URA, and compare (a) a naive MRC beamforming with (b) our proposed beamforming. The cyan diamond marks the BS position; red cross denotes the UE's estimated position used for beamforming; green cross indicates the UE's true position, from which we derive the corresponding path gain. Beam tracking animation is available as supplementary material.

To simulate the channel link quality, we designate the position of the BS to be close to the U-turn. These position estimates inform our beamforming (precoding) weight calculations. Thereafter, we use the NVidia Sionna RT library [19] for ray-tracing, enabling us to model the path loss, fading and reflections. Note that while our beamforming strategy described and simulated herein only focuses on direct LOS paths, the simulation results also models and accounts for multi-path components. An animation video of the changes in coverage map corresponding to different beamforming weights is included as supplementary material.[3] Fig. 3 provides snapshots from this simulation, comparing coverage maps at the 714th frame for the two different beamforming strategies.

To characterize the effective path gain at the actual UE's position and the impact of the position estimate, we obtain the path gain from the coverage map. This is done by averaging the values from the simulation results within a 5 m-by-5 m area centered around the UE's ground truth position. The resulting value accounts for both the impact of free-space path loss resulting from the UE's distance from the BS, as well as the directivity of power influenced by the steering direction of the beamforming weights.

Fig. 4 shows the path gains using an oracle MRC approach, assuming the true position of the UE were known, against the MRC approach based on the position mismatch (in red). This comparison clearly demonstrates instances where the communication link would be significantly degraded. However, by using our proposed beamforming scheme (weighted by the aforementioned uncertainty model assumption), we generally retain moderate path gains throughout the trajectory.

VI. Concluding Remarks

In this study, we consider a simple yet crucial scenario, underscoring the importance of beamforming optimization that

[3]https://github.com/garyleecf/odometrybeamforming

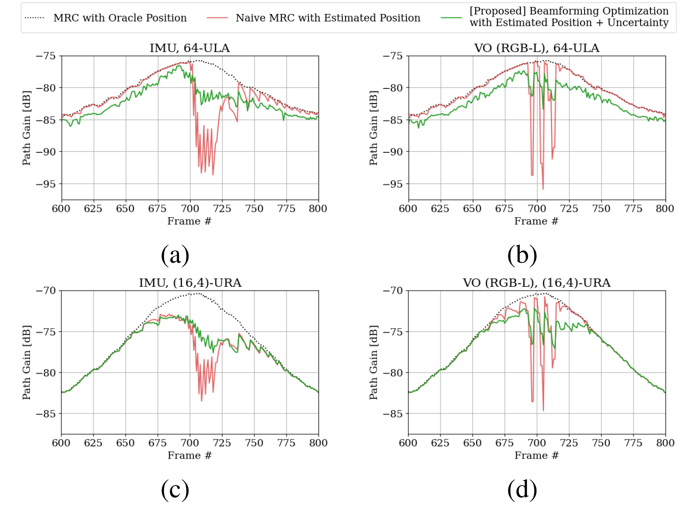


Fig. 4. Path gain simulation results along the ground-truth trajectory, using beamforming computed based on odometry-based position estimates. (a) and (b) depict the IMU and VO-based (RGB-L [18]) results on 64-ULA, and similarly, (c) and (d) depict those for a $(16, 4)$-URA.

accounts for the inherent errors in localization and odometry methods. Notably, our approach demonstrates robustness, thereby enhancing resilience against communication links dropped due to the UE's mobility. Furthermore, our proposed method reflects the advantages of overparameterization in gradient-based optimization, yielding empirical benefits. This finding could also extend to other optimization formulations, such as those discussed in [10], [11].

Moving forward, we recognize the importance of extending our methodology to more complex environments, such as NLOS conditions. Adapting our beamforming optimization (7) to incorporate a multi-path channel model is a logical next step. However, deriving accurate channel information solely from sensor measurements in such complex settings is a significant technical challenge, requiring further research.

Regarding MIMO systems, the objective function would extend beyond maximizing signal gain in intended directions, but also suppressing gains in certain AODs to minimize interference. Our proposed beamforming optimization approach may be modified to include gain penalization terms at specific AODs; nonetheless, this nuanced balance between maximizing and suppressing signal gains, particularly in the context of the uncertainties in UE and interference positioning, is a direction which we leave for future investigation.

In our simulations, we adopted a simplified statistical model to characterize position uncertainties, treating the uncertainty distribution as time-invariant across the entire trajectory. We acknowledge that a more precise characterization of these uncertainties would likely yield significantly improved results. Future research could benefit from exploring other localization and odometry algorithms, particularly those that provide finer uncertainty characterizations.

Looking ahead, integrating our beamforming approach within a comprehensive localization and mapping framework would be an important research direction, especially in characterizing the complete system solution. Such integration could enhance the reliability of wireless communication systems, especially in V2I applications, paving the way for innovative solutions for future mobile communication technology.

References

[1] M. Giordani, A. Zanella, and M. Zorzi, "Millimeter wave communication in vehicular networks: Challenges and opportunities," in *6th Int. Conf. Modern Circuits Syst. Technologies (MOCAST)*. IEEE, May. 2017, pp. 1–6.

[2] F. Jameel, S. Wyne, S. J. Nawaz, and Z. Chang, "Propagation channels for mmWave vehicular communications: State-of-the-art and future research directions," *IEEE Wirel. Commun.*, vol. 26, no. 1, pp. 144–150, 2018.

[3] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, 2017.

[4] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. De Poorter, and I. Moerman, "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Commun. Surv. Tutor.*, vol. 20, no. 4, pp. 3060–3097, 2018.

[5] M. Giordani, M. Mezzavilla, and M. Zorzi, "Initial access in 5G mmWave cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 40–47, 2016.

[6] S. A. Vorobyov, "Principles of minimum variance robust adaptive beamforming design," *Signal Processing*, vol. 93, no. 12, pp. 3264–3277, 2013.

[7] A. Klautau, N. González-Prelcic, and R. W. Heath, "LIDAR data for deep learning-based mmWave beam-selection," *IEEE Wirel. Commun. Lett.*, vol. 8, no. 3, pp. 909–912, 2019.

[8] M. Dias, A. Klautau, N. González-Prelcic, and R. W. Heath, "Position and LIDAR-aided mmWave beam selection using deep learning," in *IEEE 20th Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*. IEEE, Jul. 2019, pp. 1–5.

[9] J. Morais, A. Bchboodi, H. Pezeshki, and A. Alkhateeb, "Position-aided beam prediction in the real world: How useful GPS locations actually are?" in *IEEE Int. Conf. Commun. (ICC)*. IEEE, Jun. 2023, pp. 1824–1829.

[10] B. Flowers, X. Zhang, and S. Dey, "Low overhead codebook design for mmWave roadside units placed at smart intersections," in *IEEE 32nd Annu. Int. Symp. Pers., Indoor. Mobile Radio Commun. (PIMRC)*. IEEE, Sep. 2021, pp. 366–372.

[11] M. Alrabeiah, Y. Zhang, and A. Alkhateeb, "Neural networks based beam codebooks: Learning mmwave massive MIMO beams that adapt to deployment and hardware," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 3818–3833, 2022.

[12] C. J. Lam and A. C. Singer, "Bayesian beamforming for DOA uncertainty: theory and implementation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4435–4445, 2006.

[13] K. L. Bell, Y. Ephraim, and H. L. Van Trees, "A Bayesian approach to robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 386–398, 2000.

[14] H.-L. Song and Y.-C. Ko, "Beam alignment for high-speed UAV via angle prediction and adaptive beam coverage," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10 185–10 192, 2021.

[15] S. Arora, N. Cohen, N. Golowich, and W. Hu, "A convergence analysis of gradient descent for deep linear neural networks," *arXiv preprint arXiv:1810.02281*, 2018.

[16] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "PyTorch: An imperative style, high-performance deep learning library," *Adv. Neural Inf. Process.*, vol. 32, 2019.

[17] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," in *IEEE Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*. IEEE, Jun. 2012, pp. 3354–3361.

[18] F. Sauerbeck, B. Obermeier, M. Rudolph, and J. Betz, "RGB-L: Enhancing indirect visual SLAM using LiDAR-based dense depth maps," in *3rd Int. Conf. Comput. Control Robot. (ICCCR)*. IEEE, Mar. 2023, pp. 95–100.

[19] J. Hoydis, F. A. Aoudia, S. Cammerer, M. Nimier-David, N. Binder, G. Marcus, and A. Keller, "Sionna RT: Differentiable ray tracing for radio propagation modeling," *arXiv preprint arXiv:2303.11103*, 2023.