Ahmet M. Elbir [ID], Kumar Vijay Mishra [ID], Sergiy A. Vorobyov [ID], and Robert W. Heath Jr. [ID]

Twenty-Five Years of Advances in Beamforming

From convex and nonconvex optimization to learning techniques

©SHUTTERSTOCK.COM/TRIFF

Beamforming is a signal processing technique to steer, shape, and focus an electromagnetic (EM) wave using an array of sensors toward a desired direction. It has been used in many engineering applications, such as radar, sonar, acoustics, astronomy, seismology, medical imaging, and communications. With the advent of multiantenna technologies in, say, radar and communication, there has been a great interest in designing beamformers by exploiting convex or nonconvex optimization methods. Recently, machine learning (ML) is also leveraged for obtaining attractive solutions to more complex beamforming scenarios. This article captures the evolution of beamforming in the last 25 years from convex to nonconvex optimization and optimization to learning approaches. It provides a glimpse into these important signal processing algorithms for a variety of transmit–receive architectures, propagation zones, propagation paths, and multidisciplinary applications.

Introduction

Beamforming is ubiquitous and essential to a multitude of array processing applications, such as radar, sonar, acoustics, astronomy, seismology, ultrasound, and communications [1]. Recent advances in mobile communications, usage of large arrays, high-frequency sensors, near-field signal recovery, and smart radio environments open up interesting and novel signal processing problems in beamforming. These applications are driving the need for higher robustness, flexible deployment, and low complexity in beamforming algorithms and an emphasis on advanced signal processing that should be tailored for emerging application-specific requirements.

Early experiments with beamforming could be traced back to Guglielmo Marconi, who used a circular array with four antennas to improve the gain of trans-Atlantic Morse code transmission in 1901 [2]. A similar early demonstration of gains provided by a phased array to direct radio waves was in 1905 by Karl Ferdinand Braun, who shared the Nobel Prize in Physics with Marconi in 1909 for their contributions to wireless telegraphy [3]. In the 1940s, antenna diversity as a technique to overcome fading was developed for phased array

radars and radio astronomy [4]. By the 1950s–1960s, with the development of phased arrays for sonars, the steering of signals with antenna arrays was no longer restricted to EM waves [5].

Adaptive beamforming [6], [7] emerged in the late 1960s, wherein a processor at the antenna back end updates and compensates the array weights. In particular, Bernard Widrow introduced the least mean square algorithm to update the weights at every iteration by estimating the gradient of the mean-square error (MSE) between the desired and received signals [7]. Subsequently, J. Capon proposed selecting the weight vectors, or beamformers, to minimize the array output power. The Capon beamformer is subjected to the linear constraint that the signal of interest (SoI) does not suffer from any distortion, e.g., direction mismatch, signal fading, local scattering, etc. [6], [8]. Hence, this technique is also usually referred to as the *minimum variance* (MV) *distortionless response* (*MVDR*) beamforming.

The performance of the Capon beamformer strongly depends on the knowledge of the SoI, which is imprecise in practice because of the differences between the assumed and true array responses. The beamforming performance is usually measured by the signal-to-interference-plus-noise ratio (SINR). This may severely degrade even in the presence of small errors or mismatches in the steering vector [8]. In the past, numerous approaches were proposed to improve the robustness against errors/mismatches in the look direction [9], [10]; array manifold [11]; and local scattering [12]. These techniques were limited to only the specific mismatch they treat [13], thereby giving rise to early generalization of robust beamforming approaches, such as the sample matrix inversion (SMI) algorithm [14], robust Capon beamforming [15], eigenspace-based beamformer [16], worst case performance optimization [13], and general-rank beamformer [17], [18].

In the late 1990s and early 2000s, significant progress was made toward robust beamformer design by exploiting convex optimization [19]. These methods typically consider minimizing the effect of mismatches in the array-steering vectors and the look direction based on the worst case performance optimization [13], [15], [20]. Here, the optimization problem is cast as a second-order cone (SOC) program and efficiently solved by interior-point methods. It may also be desirable to design a robust MVDR beamformer by including the uncertainty in the array manifold via an ellipsoid or a sphere model for a particular look direction [15], [20].

During the late 2000s, certain applications of beamforming that have nonconvex objective functions or constraints gained salience. These included robust adaptive beamforming with additional constraints related to the positive semidefiniteness (PSD) of the signal covariance matrix [18], norm of the steering vectors [21], [22], [23], [24], and stochastic distortionless response [25], [26]; multicast transmit beamforming [27]; and hybrid (analog/digital) beamforming [28]. The solution to these nonconvex optimization problems usually requires recasting the problem into a tractable form through the use of, for example, semidefinite relaxation (SDR), compressed sensing (CS) [28], and alternating optimization [19]. Solving for beamforming weights is generally considered as a continuous optimization problem. However, there is a smaller body of literature [29], [30] on discrete/combinatorial techniques. Here, the beamforming weights are selected from a set of exponentials with discretized angles.

In the last decade, with the advent of new cellular communications technologies, beamforming has been extensively investigated for multiantenna systems [28]. The 4G networks (2009 to present) operating at 2.2–4.9 GHz use up to 32 antennas in a multiple-input, multiple-output (MIMO) configuration. The 5G systems (2019 to present) offer support for larger antenna arrays as well as communication at frequencies above 24 GHz. Support for larger arrays is essential in millimeter-wave (mm-wave) systems to overcome shrinking antenna sizes [31]. To reduce the hardware, cost, power, and area in mm-wave massive MIMO systems, hybrid (analog and digital) beamforming has been introduced [28], [31]. Unlike a conventional digital beamformer employing a single radio-frequency (RF) chain dedicated to each antenna, hybrid approaches employ a few (large) RF chains (analog components, e.g., phase shifters) to reduce the hardware cost. The hybrid beamformer design is also nonconvex because of the unit-modulus constraint owing to the use of phase shifters in the analog beamformers. This problem has been addressed through techniques such as sparse matrix reconstruction via CS [28], optimization over Riemannian manifolds [32], phase extraction [33], and Gram–Schmidt orthogonalization [34].

Very recently, data-driven methods, such as ML, have been leveraged to obtain beamformers. ML is a subset of artificial intelligence (AI) that allows neural networks (NNs) to learn directly from precedents, data, and examples without being explicitly programmed. Many beamformers involve nonlinear operations. In this context, NNs are particularly attractive because they successfully approximate nonlinear functions or predict the class of a function that is divided by a nonlinear decision boundary. Compared to the model-based techniques, ML has lower posttraining computational complexity, expedited design procedure, and robustness against imperfections/mismatches [35], [36], [37]. The ML-based hybrid beamforming is also envisioned as a key to realize massive MIMO architectures beyond 5G communications [38], such as 6G systems operating at terahertz (THz) bands. This is largely because ML is helpful in processing copious amounts of antenna array data generated by massive MIMO systems employed at higher frequencies.

To shed light on the evolution of beamforming techniques, this article presents an overview of the aforementioned approaches while focusing on major breakthroughs during the last 25 years. Specifically, the article aims at 1) highlighting the two significant leaps in this research, i.e.,

> **Beamforming has been used in many engineering applications, such as radar, sonar, acoustics, astronomy, seismology, medical imaging, and communications.**

convex to nonconvex optimization, and optimization- to learning-based beamforming; 2) depicting in detail the analytical background and the relevance of signal processing tools for beamforming; and 3) introducing the major challenges and emerging signal processing applications of beamforming. Figure 1 summarizes some important classes of beamformers discussed in this article.

Notation

Throughout this article, uppercase and lowercase bold letters denote matrices and vectors, respectively. Also, $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$ denote the transpose and conjugate transpose operations, respectively. For a matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ and a vector $\boldsymbol{a} \in \mathbb{C}^N$, $[\mathbf{A}]_{ij}$, $[\mathbf{A}]_k$, $\Re\{\mathbf{A}\}$ and $\Im\{\mathbf{A}\}$, and a_i correspond to the (i, j)th entry, kth column, real and imaginary parts of $\mathbf{A}$, and ith entry of $\boldsymbol{a}$, respectively, while $\mathbf{A}^\dagger$ denotes the Moore–Penrose pseudo-inverse of $\mathbf{A}$, and $\mathbf{I}$ is the identity matrix of proper size. $\|\boldsymbol{a}\|_2 = (\sum_{i=1}^N |a_i|^2)^{1/2}$ and $\|\mathbf{A}\|_{\mathcal{F}} = (\sum_{i=1}^M \sum_{j=1}^N |[\mathbf{A}]_{ij}|^2)^{1/2}$ denote the l_2 norm and Frobenius norm, respectively.

Convex optimization for beamforming

Convex optimization recasts originally difficult-to-design beamformers as computationally attractive problems that yield exact or approximate solutions through algorithms, such as interior-point methods. Its applications have traditionally advanced from the simple exact Capon approach to more complex transmit, multicast, network, and distributed beamformers; see, e.g., [19] and the references therein for details. In the following, we summarize the techniques that yield exact solutions. The approximate solutions are considered under nonconvex beamformers in the sequel.

Capon beamformer

Consider an antenna array with N elements. Define $\mathbf{a}(\theta) \in \mathbb{C}^N$ as the array response to a plane-wave narrowband SoI $s(t_i)$, $i = 1, \ldots, T$, where T is the number of snapshots arriving from the direction of arrival (DoA) angle θ. In particular, the steering vector $\mathbf{a}(\theta)$ is

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{N}} \Big[1, e^{-\mathrm{j}2\pi \frac{d}{\lambda}\sin\theta}, \ldots, e^{-\mathrm{j}2\pi \frac{(N-1)d}{\lambda}\sin\theta} \Big]^{\mathsf{T}} \tag{1}$$

where d is the element spacing, and λ is the wavelength. Then, the $N \times 1$ antenna array output is

$$\mathbf{y}(t_i) = \mathbf{a}(\theta)s(t_i) + \mathbf{e}(t_i) \tag{2}$$

where $\mathbf{e}(t_i) \in \mathbb{C}^N$ denotes the temporarily and spatially white Gaussian noise vector with variance σ^2.

The received signals are multiplied by the beamforming weights, i.e., $w_1, \ldots, w_N \in \mathbb{C}$. Therefore, the combined beamformer output becomes

$$y_o(t_i) = \mathbf{w}^{\mathsf{H}}\mathbf{y}(t_i) = \mathbf{w}^{\mathsf{H}}\mathbf{a}(\theta)s(t_i) + \mathbf{w}^{\mathsf{H}}\mathbf{e}(t_i) \tag{3}$$

where $\mathbf{w} = [w_1, \ldots, w_N]^{\mathsf{T}}$ includes the beamformer weights. To recover the signal $s(t_i)$, the beamformer weights are optimized via

$$\underset{\mathbf{w}}{\text{minimize}} \ \mathbf{w}^{\mathsf{H}}\mathbf{R}_y\mathbf{w} \quad \text{subject to} \quad \mathbf{w}^{\mathsf{H}}\mathbf{a}(\theta) = 1 \tag{4}$$
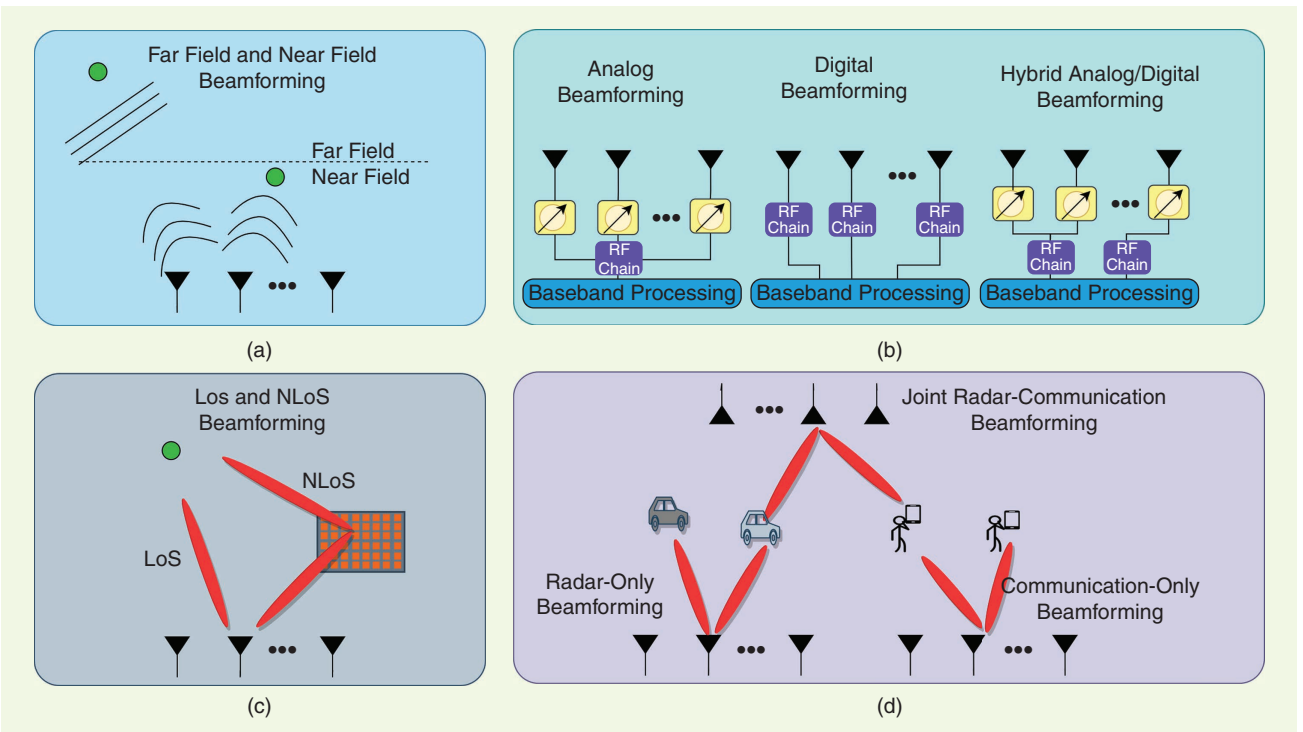
FIGURE 1. The major classes of beamforming methods by (a) transmission range: far and near fields; (b) transceiver architectures: analog, digital, and hybrid beamforming; (c) paths: LoS and NLoS beamforming, wherein the NLoS path is controlled via joint active (transmitter) and passive (intelligent reflecting surface) devices; (d) applications: radar, communications, and joint radar-communications. LoS: line-of-sight; NLoS: non-line-of-sight.

where $\mathbf{R}_y = (1/T)\Sigma_{i=1}^{T}\mathbf{y}(t_i)\mathbf{y}^{\mathsf{H}}(t_i)$ is the sample covariance matrix of the array output. The optimal solution for (4) yields the Capon beamformer [6]:

$$\mathbf{w}_{\text{opt}} = (\mathbf{a}^{\mathsf{H}}(\theta)\mathbf{R}_y^{-1}\mathbf{a}(\theta))^{-1}\mathbf{R}_y^{-1}\mathbf{a}(\theta). \qquad (5)$$

This beamformer requires the knowledge of $\mathbf{a}(\theta)$ and $\mathbf{R}_y$. Therefore, its performance depends on the accuracy of the steering vector constructed from the estimate of θ as well as the sample covariance matrix $\mathbf{R}_y$.

To stabilize the main beam response in the presence of a pointing error [9], additional constraints are added to the optimization problem as

$$\underset{\mathbf{w}}{\text{minimize}}\ \mathbf{w}^{\mathsf{H}}\mathbf{R}_y\mathbf{w} \quad \text{subject to}\ \mathbf{C}^{\mathsf{H}}\mathbf{w} = \mathbf{u} \qquad (6)$$

where L many constraints are represented by $\mathbf{C} \in \mathbb{C}^{L \times N}$ and $\mathbf{u} \in \mathbb{C}^{L}$. For example, if it is desired to maximize the beam pattern at 30° and place a null at 40°, then $\mathbf{C} = [\mathbf{a}(30°), \mathbf{a}(40°)]^{\mathsf{T}}$. and $\mathbf{u} = [1, 0]^{\mathsf{T}}$. The solution to this constrained problem is $\mathbf{w}_C = \mathbf{R}_y^{-1}\mathbf{C}(\mathbf{C}^{\mathsf{H}}\mathbf{R}_y^{-1}\mathbf{C})^{-1}\mathbf{u}$ [10].

Loaded SMI beamformer

Even in the ideal case, wherein the SoI direction θ is accurately known, beamforming performance may significantly deteriorate because of a small training sample size T. This is mitigated by adding a regularization term γ to the objective function in (4) leading, to loaded SMI (LSMI) beamforming [14]:

$$\underset{\mathbf{w}}{\text{minimize}}\ \mathbf{w}^{\mathsf{H}}\mathbf{R}_y\mathbf{w} + \gamma\|\mathbf{w}\|_2 \quad \text{subject to}\ \mathbf{w}^{\mathsf{H}}\mathbf{a}(\theta) = 1. \qquad (7)$$

Its solution is $\mathbf{w}_{\text{LSMI}} = \mathbf{R}_{\text{LSMI}}^{-1}\mathbf{a}(\theta)$, where $\mathbf{R}_{\text{LSMI}} = \mathbf{R}_y + \gamma\mathbf{I}_N$.

Robust Capon beamformer

The exact knowledge of the SoI direction θ required by the Capon beamformer is not available in practice. This is addressed by robust beamforming, which provides tolerance against the inaccuracies in the estimated SoI direction and the corresponding steering vector. A robust variant of Capon beamforming was introduced in [15], wherein the convex optimization problem is

$$\underset{\mathbf{w}}{\text{minimize}}\ \mathbf{w}^{\mathsf{H}}\mathbf{R}_y^{-1}\mathbf{w}, \quad \text{subject to}\ \|\mathbf{w} - \bar{\mathbf{a}}\|_2 \le \epsilon \qquad (8)$$

where $\bar{\mathbf{a}} = \mathbf{a}(\theta + \Delta_\theta)$ is the inaccurate steering vector for the mismatched direction $\theta + \Delta_\theta$.

Beamforming with worst case performance optimization

A more general approach is considered in [13] by taking into account the distortions in the steering vector as $\tilde{\mathbf{a}} = \mathbf{a}(\theta) + \Delta_{\mathbf{a}}$, where $\Delta_{\mathbf{a}} \in \mathbb{C}^N$ represents the steering vector distortions. As a result, the optimization problem is based on the worst case beamforming performance. Relying on the bounded Euclidean norm as $\|\Delta_{\mathbf{a}}\|_2 \le \varepsilon$ corresponding to the case of spherical uncertainty [13], the following convex problem is formulated:

$$\underset{\mathbf{w}}{\text{minimize}}\ \mathbf{w}^{\mathsf{H}}\mathbf{R}_y\mathbf{w}, \quad \text{subject to}\ |\mathbf{w}^{\mathsf{H}}\tilde{\mathbf{a}}| \ge 1, \|\Delta_{\mathbf{a}}\|_2 \le \varepsilon \qquad (9)$$

for which the LSMI-based solutions may also be obtained [8], [19]. A similar approach, called robust MV beamforming, introduced in [20], is based on ellipsoidal uncertainty. Both spherical [e.g., $\|\tilde{\mathbf{a}} - \mathbf{a}(\theta)\|_2 \le \varepsilon$ in (9)] and ellipsoidal [e.g., $(\tilde{\mathbf{a}} - \mathbf{a})^{\mathsf{H}}\mathbf{V}(\tilde{\mathbf{a}} - \mathbf{a}) \le \tilde{\varepsilon}$, where $\mathbf{V} \in \mathbb{C}^{N \times N}$ is a PSD matrix] models are used to ensure robust solutions. The latter may naturally lead to a more accurate uncertainty description [20] than that with spherical models [20], [39] if more information than just the same uncertainty radius in all mismatch dimensions is available, and an uncertainty ball is replaced by an uncertainty ellipsoid. Assuming the availability of more information about the mismatch is, however, somewhat contradictory to the notion of robustness.

The structure of the beamformer design problem also depends on the noise model. Some beamforming techniques are based on the MV criterion mentioned earlier. However, this criterion is statistically optimal only when the SoI, interference, and noise are Gaussian. The non-Gaussian case leads to a nonconvex problem as

$$\underset{\mathbf{w}}{\text{minimize}}\ \|\mathbf{Y}^{\mathsf{H}}\mathbf{w}\|_p^p, \quad \text{subject to}\ \mathbf{a}^{\mathsf{H}}(\theta)\mathbf{w} = 1 \qquad (10)$$

where $\mathbf{Y} = [\mathbf{y}(t_1), \ldots, \mathbf{y}(t_T)] \in \mathbb{C}^{N \times T}$, and $\|\mathbf{y}(t_i)\|_p^p = (\Sigma_{n=1}^{N} y_n(t_i))^{1/p}$ denotes the ℓ_p norm for $p \ge 1$. Note that (10) reduces to Capon beamforming of (4) for $p = 2$. The solution for (10) is achieved via iterative reweighted MVDR techniques [40]. In addition to generalizing the noise model, a specific choice of priors over the distribution of the beamforming weights may also be used in, say, sparsity-driven beamforming [41].

Beamforming for a general-rank source

In practice, the source signal is incoherently scattered such that the point-source assumption may not hold [17], and the array covariance matrix is no longer rank–1. Therefore, instead of a constraint on a single steering vector, the SoI covariance matrix is used. The corresponding MVDR-type optimization problem is

$$\underset{\mathbf{w}}{\text{minimize}}\ \mathbf{w}^{\mathsf{H}}\mathbf{R}_y\mathbf{w} \quad \text{subject to}\ \mathbf{w}^{\mathsf{H}}\mathbf{R}_s\mathbf{w} = 1 \qquad (11)$$

where $\mathbf{R}_s$ is the SoI covariance matrix [18]. The optimal solution to (11) is $\mathbf{w}_{\text{GR}} = \mathcal{P}[\mathbf{R}_y^{-1}\mathbf{R}_s]$, where $\mathcal{P}[\cdot]$ is the principal eigenvector operator.

Nonconvex beamformer design

Nonconvex beamformers [21], [22], [23], [24], [25], [26], [27], [28], [42] tackle the design problem by recasting or relaxing it into tractable convex forms. This may be achieved by dropping the nonconvex constraints or decoupling the beamforming design into multiple convex subproblems.

PSD-constrained beamforming

The general-rank beamforming solution in (11) requires the knowledge of signal covariance matrix $\mathbf{R}_s$, which is not always available [17], [18]. The actual signal correlation matrix is, then, not guaranteed to be PSD and usually modeled as $\tilde{\mathbf{R}}_s = \mathbf{R}_s + \Delta_s$. To guarantee the PSD-ness of $\tilde{\mathbf{R}}_s$ decompose

it as $\tilde{\mathbf{R}}_s = \mathbf{Q}\mathbf{Q}^H$ with the mismatch parameter $\Delta_{\mathbf{Q}}$ bounded as $\|\Delta_Q\|_2 \le \varepsilon_Q$. The resulting nonconvex problem is

$$\begin{aligned}\underset{\mathbf{w}}{\text{minimize}}\ &\max_{\|\Delta_y\|_2 \le \varepsilon_y} \mathbf{w}^H(\mathbf{R}_y + \Delta_y)\mathbf{w} \\ \text{subject to}\ &\min_{\|\Delta_Q\|_2 \le \varepsilon_Q} \mathbf{w}^H(\mathbf{Q}+\Delta_Q)^H(\mathbf{Q}+\Delta_Q)\mathbf{w} \ge 1 \end{aligned} \quad (12)$$

where Δ_y, with $\|\Delta_y\|_2 \le \varepsilon_y$, represents the mismatch in $\mathbf{R}_y$. The efficient solution to the nonconvex problem in (12) is obtained via the polynomial-time difference-of-convex functions algorithm [18].

Norm-constrained beamforming based on steering vector estimation

Apart from the uncertainty constraint (8) of the robust Capon beamformer [15], [21] considers an additional norm constraint for beamformer weights in a more general setting as

$$\underset{\mathbf{w}}{\text{minimize}}\ \mathbf{w}^H\hat{\mathbf{R}}^{-1}\mathbf{w} \quad \text{subject to}\ \|\mathbf{a} - \tilde{\mathbf{a}}\|_2 \le \epsilon_a, \|\mathbf{a}\|_2^2 = N \quad (13)$$

which is identical to (8) and convex without the constraint $\|\mathbf{a}\|_2^2 = N$. The nonconvex problem in (13) is called *doubly constrained* robust Capon beamforming [21]. It is iteratively solved by interpreting the optimization as a covariance fitting problem. Thus, a robust beamformer is obtained by robustly estimating the array-steering vector. This formulation was further improved in [23], where the difference between the actual and presumed steering vectors is iteratively estimated without making any assumption on either the norm of the mismatch vector or its probability distribution.

The solution developed in [23] has led to a formulation in [24] of a new constraint, which guarantees that an estimate of the source steering vector does not converge to any steering vectors of interference signals as well as their linear combinations. This steering vector estimation problem is

$$\underset{\hat{\mathbf{a}}}{\text{minimize}}\ \hat{\mathbf{a}}^H\hat{\mathbf{R}}^{-1}\hat{\mathbf{a}} \quad \text{subject to}\ \|\hat{\mathbf{a}}\|_2^2 = N,\ \hat{\mathbf{a}}^H\tilde{\mathbf{C}}\hat{\mathbf{a}} \le \Delta_0 \quad (14)$$

where the last constraint is new; $\hat{\mathbf{a}} \in \mathbb{C}^N$ is the estimate of $\mathbf{a}$; $\tilde{\mathbf{C}} = \int_{\tilde{\Theta}} \mathbf{a}(\theta)\mathbf{a}^H(\theta) d\theta$ $\tilde{\Theta}$ is the complement of the angular sector $\Theta = [\theta_{\min}, \theta_{\max}]$ where the desired signal is located; and Δ_0 is a uniquely selected value for a given Θ, that is, $\Delta_0 \triangleq \max_{\theta \in \Theta} \mathbf{a}^H(\theta)\tilde{\mathbf{C}}\mathbf{a}(\theta)$, representing the boundary line to distinguish approximately whether or not the direction of $\mathbf{a}$ is in the actual signal angular sector Θ.

To account for gain perturbations in the steering vector, [22] added the double-sided norm constraint to the problem (14) as

$$\begin{aligned}\underset{\hat{\mathbf{a}}}{\text{minimize}}\ &\hat{\mathbf{a}}^H\hat{\mathbf{R}}^{-1}\hat{\mathbf{a}} \quad \text{subject to}\ \hat{\mathbf{a}}^H\mathbf{C}\hat{\mathbf{a}} \ge \Delta_1, \\ &N(1-\eta_1) \le \|\hat{\mathbf{a}}\|_2^2 \le N(1+\eta_2),\ \left\|\mathbf{V}^H(\hat{\mathbf{a}}-\mathbf{a}_0)\right\|_2^2 \le \epsilon_u \end{aligned} \quad (15)$$

where $\mathbf{a}_0 = \mathbf{a}(\theta_0)$, $\theta_0 = (\theta_{\max} + \theta_{\min})/2$ is the middle value of the region Θ; $\mathbf{V} \in \mathbb{C}^{N \times N}$ denotes a generalized similarity constraint together with $\mathbf{a}_0$ and ϵ_u; $\mathbf{C} = \int_\Theta \mathbf{a}(\theta)\mathbf{a}^H(\theta) d\theta$; and Δ_1,

η_1, and η_2 are selected values. In (15), the generalized similarity condition implies that imperfect knowledge of the desired steering vector $\hat{\mathbf{a}}$ is described as in a convex set (in particular, an ellipsoidal set when $\mathbf{V}$ is of full row rank).

All of these problems are nonconvex but can be often exactly solved through SDR, iterative SOC program, quadratic matrix inequality, and bilinear matrix inequality approaches.

Chance-constrained beamforming

In many applications, it is more natural that the distortionless constraint is satisfied with a certain probability. This leads to the chance-constrained robust adaptive beamforming problem [25]:

$$\underset{\mathbf{w}}{\text{minimize}}\ \mathbf{w}^H\hat{\mathbf{R}}\mathbf{w} \quad \text{subject to}\ \Pr\left\{\left|\mathbf{w}^H\tilde{\mathbf{a}}\right| \ge 1\right\} \ge p \quad (16)$$

where p is a certain preselected probability value, and $\Pr\{\cdot\}$ stands for the probability operator. This problem corresponds to minimizing the beamformer output power subject to the stochastic constraint that the probability of the signal distortionless response is greater than or equal to some selected value p. The constraint may also be viewed as a nonoutage probability constraint where the outage probability $p_{\text{out}} = 1 - p$ is defined as that of violating the inequality $\left|\mathbf{w}^H\tilde{\mathbf{a}}\right| \ge 1$ for a random $\tilde{\mathbf{a}}$ that consists of a presumptive steering vector and the mismatch that is assumed to be random. Problem (16) is nonconvex and specified by the mismatch distribution. The solutions of (16) for the case of Gaussian-distributed mismatch of the signal steering vector and for the worst case distribution are well approximated by the corresponding SOC programs [25].

In [26], a chance-constrained nonconvex formulation of robust adaptive beamforming considers a more practical scenario, wherein both interference-plus-noise covariance matrix $\mathbf{R}_{i+n}$ and the true steering vector $\mathbf{a}$ are not precisely known. It also shows the chance-constrained beamformer to have a higher output SINR than other convex (LSMI) and nonconvex (worst case optimization) beamformers [26]. Considering both $\mathbf{R}_{i+n}$ and $\mathbf{a}$ as random variables, the robust adaptive beamforming becomes

$$\begin{aligned}\underset{\mathbf{w}}{\text{minimize}}\ &\max_{G_1 \in \mathcal{S}_1} E_{G_1}\{\mathbf{w}^H\mathbf{R}_{i+n}\mathbf{w}\} \\ \text{subject to}\ &\min_{G_2 \in \mathcal{S}_2} E_{G_2}\{\mathbf{w}^H\mathbf{a}\mathbf{a}^H\mathbf{w}\} \ge 1 \end{aligned} \quad (17)$$

where $E_{G_1}\{\cdot\}$ $(E_{G_2}\{\cdot\})$ denotes the statistical expectation under the distribution G_1 (G_2), and $\mathcal{S}_1$ $(\mathcal{S}_2)$ is a set of distributions G_1 (G_2) for random matrix $\mathbf{R}_{i+n}$ (random vector $\mathbf{a}$) as, respectively,

$$\mathcal{S}_1 = \left\{ G_1 \in \mathcal{M}_1 \left| \begin{array}{l} \Pr_{G_1}\{\mathbf{R}_{i+n} \in \mathcal{Z}_1\} = 1 \\ E_{G_1}\{\mathbf{R}_{i+n}\} \succeq \mathbf{0} \\ \left\|E_{G_1}\{\mathbf{R}_{i+n}\} - \mathbf{S}_0\right\|_{\mathcal{F}} \le \rho_1 \end{array} \right. \right\} \quad (18)$$

and

$$\mathcal{S}_2 = \left\{ G_2 \in \mathcal{M}_2 \left| \begin{array}{l} \Pr_{G_2}\{\mathbf{a} \in \mathcal{Z}_2\} = 1 \\ E_{G_2}\{\mathbf{a}\} = \mathbf{a}_0 \\ E_{G_2}\{\mathbf{a}\mathbf{a}^H\} = \mathbf{\Sigma} + \mathbf{a}_0\mathbf{a}_0^H \end{array} \right. \right\} \quad (19)$$

where $\mathcal{M}_1$ and $\mathcal{M}_2$ are sets of all probability measures; $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are Borel sets; $\mathbf{S}_0$ is the empirical mean of $\mathbf{R}_{i+n}$, that is, the sample covariance matrix $\mathbf{R}_y$; and $\text{Pr}_{G_1}\{\cdot\}$ is the probability of an event under the distribution G_1. Assume the mean $\mathbf{a}_0$ and covariance matrix $\boldsymbol{\Sigma} \succ \mathbf{0}$ of random vector $\mathbf{a}$ under the true distribution $\bar{G}_2$ are known. Then, the set $\mathcal{S}_2$ includes all probability distributions on $\mathcal{Z}_2$ that have the same first- and second-order moments as $\bar{G}_2$. This problem is called *distributionally robust beamforming* because it considers distributional uncertainty in both the steering vector and $\mathbf{R}_{i+n}$.

Multicast transmit beamforming

In wireless communications, multicast beamforming is used for broadcasting data streams $s(t_i)$ toward multiple radio receivers. Consider a transmitter with an N-element antenna array that aims to deliver a signal to U single-antenna users. Denote the wireless channel between the transmitter and the uth receiver by $\mathbf{h}_u \in \mathbb{C}^N$. Then, for the beamformed transmitted signal $\mathbf{x}(t_i) = \mathbf{w}s(t_i)$, the received signal at the uth user is $y_u(t_i) = \mathbf{h}_u^H \mathbf{x}(t_i) + e_u(t_i)$, where $e_u(t_i)$ is the noise signal with variance σ_u^2. Then, the multicast beamforming problem is [27]

$$\underset{\mathbf{w}}{\text{minimize}} \|\mathbf{w}\|_2 \quad \text{subject to} \quad \left| \mathbf{w}^H \tilde{\mathbf{h}}_u \right| \geq 1, \quad u \in \{1, \ldots, U\} \quad (20)$$

where $\tilde{\mathbf{h}}_u = \mathbf{h}/\sqrt{\rho_{\min,u} \sigma_u^2}$ is the normalized channel vector with the minimum received signal-to-noise ratio (SNR) $\rho_{\min,u}$

and the noise variance σ_u^2 for the uth receiver. The optimization in (20) is a quadratically constrained quadratic programming problem with nonconvex constraints. A rigorous solution is based on reformulating the problem using SDR. To this end, define an $N \times N$ rank-one matrix $\mathbf{M} = \mathbf{w}\mathbf{w}^H$. Then, the rank constraint is removed to recast the problem in a convex form as

$$\underset{\mathbf{M}}{\text{minimize}} \, \text{trace}\{\mathbf{M}\} \quad \text{subject to} \quad \text{trace}\{\mathbf{M}\mathbf{D}_u\} \geq 1, \mathbf{M} \succeq \mathbf{0} \quad (21)$$

where $\mathbf{D}_u = \tilde{\mathbf{h}}_u \tilde{\mathbf{h}}_u^H$, and the beamformer weight is obtained via eigenvalue decomposition of $\mathbf{M}$. A more accurate solution to (20) is obtained by rewriting $\mathbf{M} = \mathbf{w}_1 \mathbf{w}_2^H$ and then alternatingly solving for $\mathbf{w}_1$ and $\mathbf{w}_2$ using an iterative procedure until convergence [30].

Hybrid analog/digital beamforming

Compared to analog- and digital-only beamformers, hybrid analog/digital beamforming architecture may have a lower hardware cost while also providing satisfactory spectral efficiency (SE) and multiple beams (Figure 2). In fact, for massive antenna array processing applications, such as 5G communications, hybrid beamforming has emerged as the preferred means to realize large arrays with only a moderate increase in baseband signal processing [31], [33].

Consider a hybrid beamforming scenario, wherein the transmitter employs N antennas and N_{RF} RF chains to send N_S
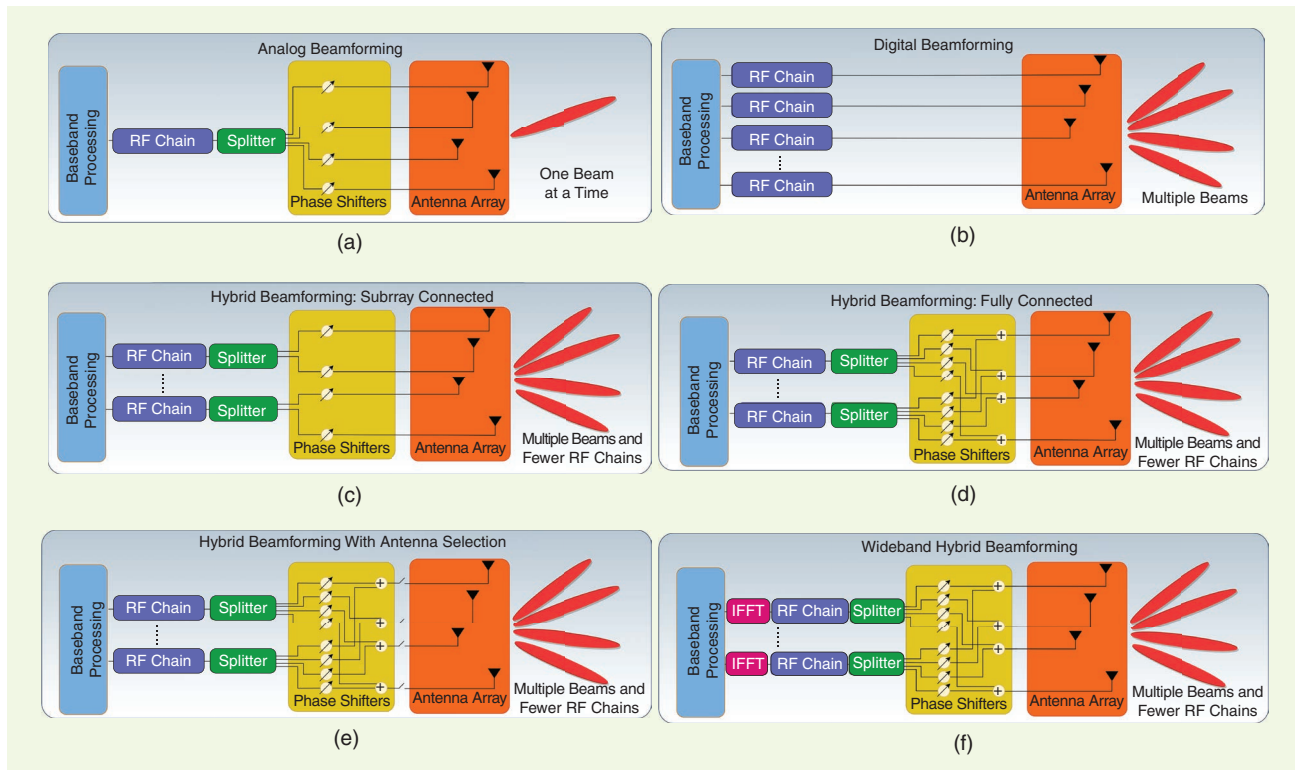
FIGURE 2. The transmitter architectures for (a) analog, (b) digital, and (c)–(f) hybrid beamforming. Analog beamforming generates only one beam because it employs a single RF chain. On the other hand, multiple beams are obtained via digital beamformers but at the cost of multiple RF chains. It is possible to generate multiple beams with fewer RF chains in the hybrid approach through configurations such as (c) subarray connected, (d) fully connected, (e) sparse antenna-selective, and (f) wideband architectures. IFFT: inverse fast Fourier transform.

data streams. Denote the analog and digital beamformers by matrices $\mathbf{F}_{\mathrm{RF}} \in \mathbb{C}^{N \times N_{\mathrm{RF}}}$ and $\mathbf{F}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{RF}} \times N_{\mathrm{S}}}$, respectively. Here, each element of $\mathbf{F}_{\mathrm{RF}}$ has a constant modulus because they are realized by phase shifters, i.e., $[\mathbf{F}_{\mathrm{RF}}]_{i,j} = 1/\sqrt{N}$ for $i = 1, \ldots, N$, $j = 1, \ldots, N_{\mathrm{RF}}$. The transmitted signal is $\mathbf{x} = \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{s}$. The goal is to maximize mutual information

$$\mathcal{I}(\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}) = \log_2 \det\left(\mathbf{I}_{N_S} + \frac{\kappa}{N_{\mathrm{S}} \sigma_n^2} \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^{\mathsf{H}} \mathbf{F}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{H}^{\mathsf{H}}\right)$$

where $\mathbf{H} \in \mathbb{C}^{N \times N_{\mathrm{R}}}$ is the wireless channel matrix, N_{R} is the number of antennas at the receiver, κ is the average received power, and σ_n^2 is the noise power [28]. The hybrid beamforming problem is

$$\underset{\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}}{\text{maximize}} \ \mathcal{I}(\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}})$$
$$\text{subject to} \ \|\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}\|_{\mathcal{F}} = N_{\mathrm{S}}, |[\mathbf{F}_{\mathrm{RF}}]_{i,j}| = \frac{1}{\sqrt{N}} \quad (22)$$

which is nonconvex because of the constant-modulus constraint. The product $\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}$ also makes this problem nonlinear. Recast (22) to an equivalent form by minimizing the Euclidean cost between the hybrid beamformer $\mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}}$ and the unconstrained baseband-only beamformer $\mathbf{F}_{\mathrm{C}} \in \mathbb{C}^{N \times N_{\mathrm{S}}}$ as

$$\underset{\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}}{\text{minimize}} \ \|\mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} - \mathbf{F}_{\mathrm{C}}\|_{\mathcal{F}}$$
$$\text{subject to} \ \|\mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}}\|_{\mathcal{F}} = N_{\mathrm{S}}, |[\mathbf{F}_{\mathrm{RF}}]_{i,j}| = \frac{1}{\sqrt{N}} \quad (23)$$

where $\mathbf{F}_{\mathrm{C}}$ is obtained from singular value decomposition of the channel matrix $\mathbf{H}$ [31]. In the wideband scenario, subcarrier-dependent (SD) digital beamformers are used, and the resulting signal is transformed to the time domain via the inverse fast Fourier transform (Figure 2). Then, subcarrier-independent analog beamformers are employed for all subcarriers because the direction of the generated beam does not change significantly with respect to subcarriers in the mm-wave band [31], [43]. The hybrid beamforming problem for a wideband system with M subcarriers is

$$\underset{\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}[m]}{\text{minimize}} \ \|\mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}}[m] - \mathbf{F}_{\mathrm{C}}[m]\|_{\mathcal{F}}$$
$$\text{subject to} \ \|\mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}}[m]\|_{\mathcal{F}} = M N_{\mathrm{S}}, |[\mathbf{F}_{\mathrm{RF}}]_{i,j}| = \frac{1}{\sqrt{N}} \quad (24)$$

where $\mathbf{F}_{\mathrm{BB}}[m]$ is the SD digital beamformer that corresponds to the mth subcarrier, $m \in \mathcal{M} = \{1, \ldots, M\}$.

For the nonconvex hybrid beamforming formulated in (23), the traditional route is to alternately optimize each ($\mathbf{F}_{\mathrm{RF}}$ and $\mathbf{F}_{\mathrm{BB}}$) beamformer iteratively while keeping the other one fixed [28], [32], [33]. This has been shown to provide satisfactory SE performance, often close to that of digital-only beamformers, i.e., $\mathbf{F}_{\mathrm{C}}$ [28], [32]. During these alternations, while estimation of digital beamformer $\mathbf{F}_{\mathrm{BB}}$ is straightforward as $\mathbf{F}_{\mathrm{BB}} = \mathbf{F}_{\mathrm{RF}}^{\dagger} \mathbf{F}_{\mathrm{C}}$, the analog beamformer $\mathbf{F}_{\mathrm{RF}}$ is difficult to obtain. Often $\mathbf{F}_{\mathrm{RF}}$ is obtained in terms of the steering vectors via CS-based techniques, e.g., orthogo-

nal matching pursuit (OMP). Here, a dictionary of possible steering vectors or atoms is employed, and the beamformers are iteratively selected from these atoms based on the similarity between the dictionary and the measurements (i.e., channel data) [28]. In manifold optimization (MO)-based approaches [32], the search space of $\mathbf{F}_{\mathrm{RF}}$ is regarded as a Riemannian submanifold of $\mathbb{C}^N$ with a complex circle manifold to account for the constant-modulus constraint. Then, the analog and digital beamformers are alternatingly optimized. This method aims to solve the unconstrained optimization problem $\min_{\mathbf{x}} f(\mathbf{x})$, $\mathbf{x} \in \mathbb{C}^n$ where $f(\mathbf{x})$ is the cost function, and vector $\mathbf{x} = \mathrm{vec}(\mathbf{F}_{\mathrm{RF}})$. To ensure global convergence, the cost function is defined over the Riemannian manifold $\mathcal{M} = \{\mathbf{x} \in \mathbb{C}^N | x_n^* x_n = 1, n = 1, \ldots, N\}$. Then, $\mathbf{x}$ is iteratively computed and the solution becomes $\mathbf{x}_{k+1} = \mathrm{Retr}_{\mathbf{x}_k}(-\alpha_k \mathrm{grad} \, f(\mathbf{x}_k))$, where Retr is the retraction on $\mathcal{M}$, and $\mathrm{grad} \, f(\mathbf{x}_k)$ denotes the Riemannian gradient [32].

The implementation of hybrid analog/digital beamforming imposes another constraint in the system design: a limited number of phase shifters and analog-to-digital converters (ADCs). Although the power consumption of phase shifters is typically lower than that of baseband beamformers, their number increases with the number of antennas. The implementation of hybrid analog/digital beamformers becomes more complex and expensive at higher frequencies (e.g., the upper mm-wave and THz). As an alternative, lens-based beamformers have been proposed [44]. Instead of using a phase shifter network, they use lenses to generate a directional beam from the EM sources placed at the focal points of the lenses. Thus, lens-based beamformers offer reduced computational complexity when compared with phase shifter-based architectures. Lens-based beamformers, though, only realize directional beams and not more sophisticated beam patterns, as may be useful in a spatial multiplexing or interference cancellation setting. A low-power design in [45] suggests using Butler matrices, which consist of an $N \times N$ matrix of hybrid couplers and fixed phase shifters.

Low-resolution ADCs

Low-resolution (1–3-bit) ADCs for digital beamformers bring down the overall power consumption and hardware cost. In particular, 1-bit ADCs do not require hardware components, such as automatic gain control and linear amplifiers. Hence, the corresponding RF chain is implemented cost-efficiently [46]. Denote the received signal at the receiver and the corresponding beamformer matrix to be $\mathbf{r} \in \mathbb{C}^{N_R}$ and $\mathbf{W}_{\mathrm{RF}} \in \mathbb{C}^{N_R \times N_S}$, respectively. Then, the received signal sampled by low-resolution ADCs is $\mathbf{r}_q = Q_b(\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{r})$, where $Q_b(\cdot)$ is the quantization operator with b-bit resolution. The received signal $\mathbf{r}_q$ is then used to design the receiver via zero-forcing or maximum-rate-combining techniques [42], [46].

Finite-resolution phase shifters

In practice, continuous-valued phase angles are expensive to implement, and finite-resolution phase shifters may be used

with low-resolution ADCs. Here, the beamformer weights are selected from the finite set $\mathcal{W} = \{1, \omega, \omega^2, \ldots, \omega^{2^b-1}\}$, where

$$\omega = \frac{1}{\sqrt{N}} e^{\frac{j2\pi}{2^b}}$$

and b is the number of bits. Then, the constant-modulus constraint in (23) is replaced by $[\mathbf{F}_{\mathrm{RF}}]_{i,j} \in \mathcal{W}$. A feasible solution to hybrid beamforming with finite resolution is to first solve (23) under the infinite resolution assumption and then quantize the phase elements of the beamformers [33].

Figure 3(a) shows the comparison of fully digital beamforming and hybrid beamforming with low-resolution phase shifters. The hybrid architecture with MO-based design has a performance very close to that of fully digital beamformers. The OMP with $b = 5$-bit phase shifters performs closest to infinite-resolution phase shifters. The gap from the fully digital performance is larger for OMP-based techniques compared to MO-based beamforming.

Learning-based beamforming

Lately, as has been the case with many signal processing problems, beamforming has also not remained untouched by ML techniques. In learning-based hybrid beamforming, the problem is approached from a model-free viewpoint by constructing a nonlinear mapping between the input data (e.g., the channel matrix and array output) and output (beamformers) of a learning model [35], [36], [37]. This method has the following advantages over model-based techniques:

■ The model-free/data-driven structure of a learning-based approach yields a robust performance in terms of SE against the corruptions (e.g., a mismatched number of received paths or imperfectly estimated channel gain and path directions [36], [37]) in the input.

■ Learning techniques extract feature patterns in the data. Hence, they easily update incoming/future data and adapt in response to environmental changes. The model-based beamformers lack these abilities and may employ statistical predictive algorithms [see Figure 3(c)].
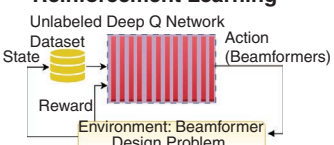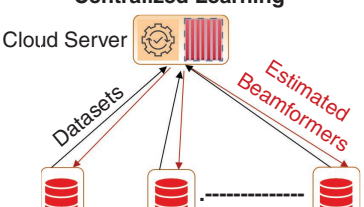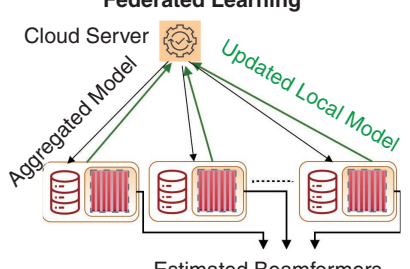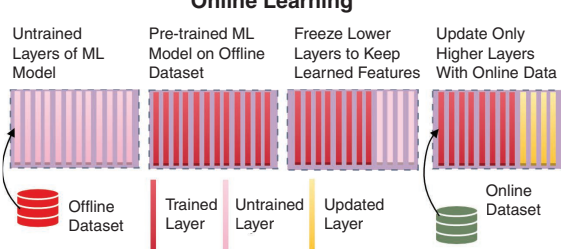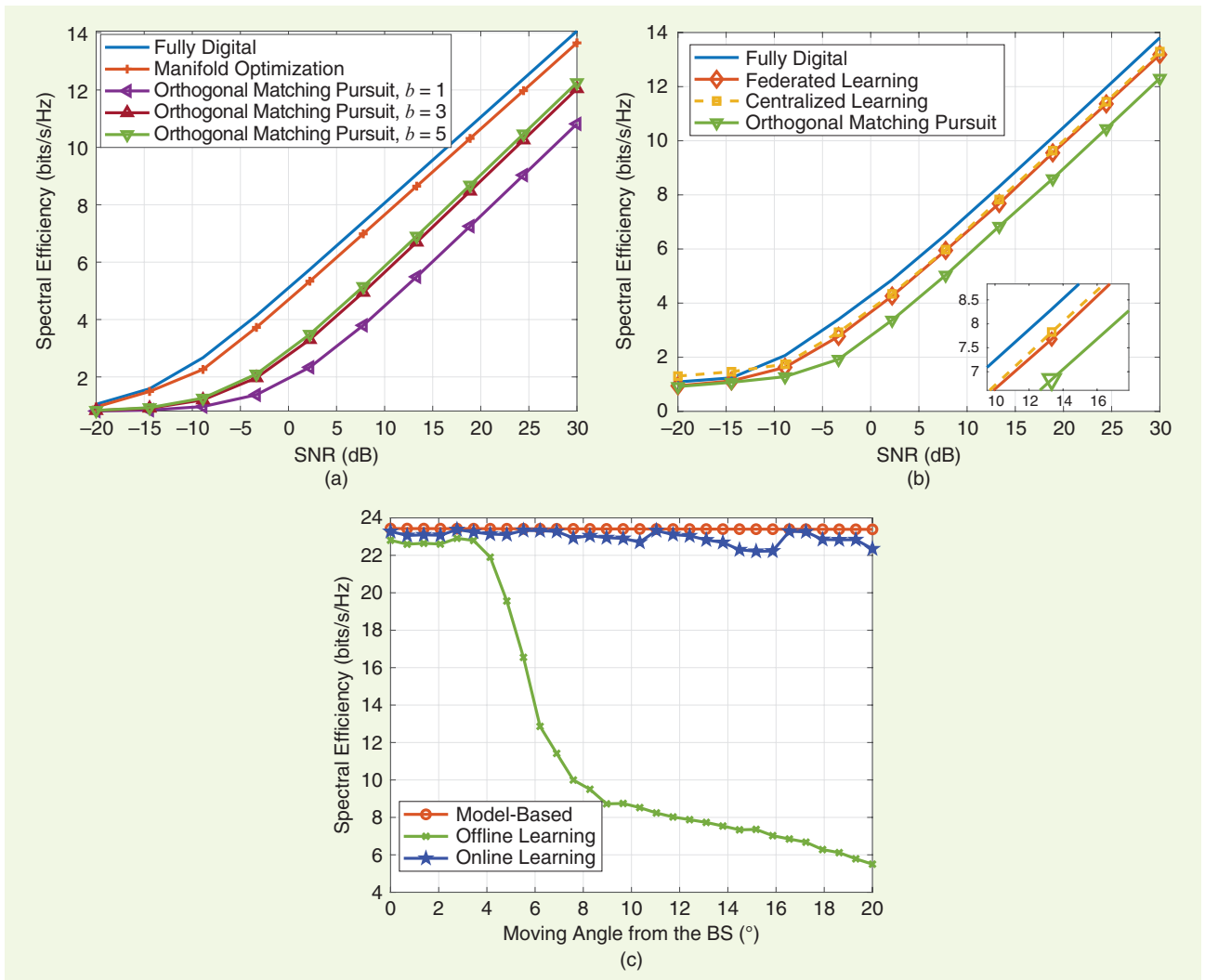

FIGURE 3. The SE performance of various hybrid beamforming approaches: (a) low-resolution phase shifters, (b) learning-based and model-based techniques, and (c) offline and online learning. Here, the channel is realized with three paths, the number of BS antenna elements $N = 100$, the number of users $U = 8$, and the number of user antennas $N_{\mathrm{R}} = 16$.

- Learning exhibits lower computational complexity in the prediction stage than optimization.

Through parallel processing, ML significantly (~10-fold [36]) reduces the computational times. On the other hand, a parallel implementation of conventional convex/nonconvex optimization-based beamforming is not straightforward. Beginning from the earlier simpler networks, such as mul-tilayer perception, to more complex deep learning models like convolutional NNs (CNNs), ML has come a long way in successfully performing feature extraction for analog and digital beamformers [47]. Table 1 summarizes various learning models, including the well-known unsupervised/supervised learning (UL/SL) and the more recent federated learning (FL).

Table 1. Learning models.

Network Model	Data	Application in Beamforming
Unsupervised Learning	Unlabeled	*Fast beamforming*: Minimize a given optimization objective to implicitly obtain the beamformers. Unsupervised learning is useful, especially for mobile transmitters, where labels are not available.
Supervised Learning	Labeled	*Uplink/downlink beamforming*: The network is trained to construct a nonlinear relationship between the input and the labeled data (beamformers).
Reinforcement Learning	Unlabeled	*Uplink/downlink beamforming*: The network learns the beamformers based on a reward/punishment mechanism in accordance with optimizing the overall system's SE.
Centralized Learning	Labeled/ unlabeled	*Uplink multiuser beamforming*: The training datasets are transmitted to a centralized cloud server, wherein the model is trained. Posttraining, each user sends the input data (channel) to the server that sends the output (estimated beamformers) to the users.
Federated Learning	Labeled/ unlabeled	*Downlink multiuser beamforming*: Instead of transmitting the whole dataset to the cloud server, each user processes its own local dataset, computes the corresponding model update, and transmits only the updates to the server. Then, the server broadcasts the aggregated model updates to the users, which can estimate their own beamformers.
Online Learning	Labeled	*Adaptive beamforming*: The learning model is updated when the prediction performance degrades because of deviations in the input compared to the training data.

UL, SL, and semi-supervised learning

UL studies the clustering of unlabeled data into smaller sets by exploiting the hidden features/patterns derived from the dataset, for which an answer key (label) is not provided beforehand. Hence, the "distance" between the training data samples is optimized without prior knowledge of the "meaning" of each clustered set. In SL, however, the labeled data are used for model training while minimizing the error between the label and the model's response. The cost function of the training is generally the MSE, but other functions (e.g., the mean error, mean absolute error, cross entropy, and Kullback–Leibler divergence) may also be used. Note that beamforming may be cast as either a regression (the output is the beamformer weights) or a classification (the output is an index of a vector from a predefined set of possible beamformers) problem. SL is widely used for several applications of beamformer design in radar and communications [43].

Define $\mathcal{X} \in \mathbb{R}^{N_{in}}$ and $\mathcal{Y} \in \mathbb{R}^{N_{out}}$ as the input and label data of a learning model whose real-valued learnable parameters are stacked into the vector $\boldsymbol{\Theta} \in \mathbb{R}^Q$. Then, the relationship between the input $\mathcal{X} \in \mathbb{R}^{N_{in}}$ and output $\mathcal{Y} \in \mathbb{R}^{N_{out}}$ is represented by a nonlinear function $f(\boldsymbol{\Theta}, \mathcal{X}) : \mathbb{R}^{N_{in}} \to \mathbb{R}^{N_{out}}$ such that $\mathcal{Y} = f(\mathcal{X}|\boldsymbol{\Theta})$ The input data are, say, the vectorized elements of the channel matrix $\mathbf{H}$ as $\mathcal{X} = [\text{vec}\{\Re\{\mathbf{H}\}^\top\}, \text{vec}\{\Im\{\mathbf{H}\}^\top\}]^\top$, and the labels are beamformers. In the case of the unit-modulus constraint, it suffices to represent the beamformers in terms of only the angle, i.e., $\mathcal{Y} = \angle\{\mathbf{F}_{RF}\}$. Note that the baseband beamformers are readily computed as $\mathbf{F}_{BB} = \mathbf{F}_{RF}^\dagger \mathbf{F}_C$ [28].

Apart from hybrid beamforming, ML techniques have been applied to other applications, such as robust beamformers [35]. Here, the sample covariance matrix is fed to a CNN whose output is the beamformer weights. The labels are obtained by solving the robust Capon beamformer problem in (8). The training dataset was $\mathcal{D} = \{\mathcal{D}_1, \ldots, \mathcal{D}_J\}$, where $\mathcal{D}_i = (\mathcal{X}_i, \mathcal{Y}_i)$ denotes the ith input–output sample for $i = 1, \ldots, J$. The model is trained by minimizing the MSE cost

$$\frac{1}{J} \sum_{i=1}^{J} \left\| \mathcal{Y}_i - f(\mathcal{X}_i|\boldsymbol{\Theta}) \right\|_2^2$$

over $\boldsymbol{\Theta}$. Posttraining, the learned parameters are used for prediction purposes for beamforming.

The acoustic beamformers in [48] are obtained via semi-supervised learning (SSL), where both labeled and unlabeled data are used. When a small set of labeled data are available in addition to a large volume of unlabeled data, using both sets in SSL is more advantageous than SL alone.

Reinforcement learning

In reinforcement learning (RL), the learning model is initialized from a random state, and the algorithms learn to react to the channel conditions on their own [49]. The model accepts the analog and baseband beamformers of the previous state as input and then updates the model parameters by taking into account the corresponding average rate as a reward. In general, RL has autonomous AI agents that gather their own data and improve based on their trial-and-error interactions with the environment. It shows a lot of promise in basic research. However, so far, RL has been harder to use in real-world beamformer applications because its dataset does not include labels. Consequently, RL requires longer training times for learning the features of wireless channels, especially in dynamic, short-coherence time scenarios.

> Using the same NN structures, CL has a better performance than FL because the former has access to the whole dataset at once, whereas the latter employs decentralized training.

Online learning

The online learning (OL) algorithm involves a learning model whose parameters are updated when there is a significant change in the received input data. For example, consider the beamformer design for a wireless communications system [Figure 3(c)], wherein the user is moving away in the DoA domain from the base station (BS). Then, the received array data become significantly different from the collected offline training data, thereby degrading the network performance. Here, hybrid beamforming and channel estimation may be performed jointly because the beamformer weights are directly related to the channel matrix. Moreover, OL is a suitable choice for this problem [36]; it updates the model parameters when the normalized MSE of channel estimates is higher than a predetermined threshold. From Figure 3(c), the learning model requires retraining every $\sim 4°$ for a massive MIMO scenario.

FL

Compared to centralized learning (CL), FL is more suited for multiuser scenarios. Using the same NN structures, CL has a better performance than FL because the former has access to the whole dataset at once, whereas the latter employs decentralized training. The FL is ideal for downlink, wherein the trained model is available to the user at the network edge. As an example, consider a downlink scenario wherein U communications users collaborate to train a model with learnable parameters $\boldsymbol{\Theta}$ with local datasets $\mathcal{D}^{(u)} = (\mathcal{X}^{(u)}, \mathcal{Y}^{(u)})$ for $u = 1, \ldots, U$. Here, the output data $\mathcal{Y}^{(u)}$ are the beamformer weights corresponding to the uth user. The FL-based training problem minimizes the averaged local cost

$$\min_{\boldsymbol{\Theta}} \frac{1}{U} \sum_{u=1}^{U} \mathcal{L}_u(\boldsymbol{\Theta})$$

where $i = 1, \ldots, J_u$ and $J_u = |\mathcal{D}^{(u)}|$ denotes the number of samples in $\mathcal{D}^{(u)}$, over $\boldsymbol{\Theta}$. Different than the cost in the "UL, SL, and Semi-supervised Learning" section, the local cost here is

$$\mathcal{L}_u(\boldsymbol{\Theta}) = \frac{1}{J_u} \sum_{i=1}^{J_u} \left\| f(\mathcal{X}_i^{(u)}|\boldsymbol{\Theta}) - \mathcal{Y}_i^{(u)} \right\|_2^2$$

for the uth user. This is efficiently solved by iteratively applying gradient descent, which updates the model parameter at the tth iteration as

$$\boldsymbol{\Theta}_{t+1} = \boldsymbol{\Theta}_t - \eta \frac{1}{U} \sum_{u=1}^{U} \boldsymbol{\beta}_u(\boldsymbol{\Theta}_t)$$

where $\boldsymbol{\Theta}_t$ is the computed model parameter vector at iteration t, $\boldsymbol{\beta}_u(\boldsymbol{\Theta}_t) = \nabla \mathcal{L}_u(\boldsymbol{\Theta}_t) \in \mathbb{R}^Q$ is the gradient vector, and η is the learning rate. Figure 3(b) compares the performance of FL and CL with model-based techniques, such as OMP, and the fully digital beamformer in terms of SE [50]. Both CL and FL outperform OMP, but the performance gap between CL and FL increases with the nonuniformity of the local dataset.

Emerging applications

Research in beamforming continues to be highly active in light of emerging applications and theoretical advances. For example, the hybrid approach of a model-driven network or deep unfolding for beamforming [51] allows for bounding the complexity of algorithms while also retaining their performance. Convolutional beamformers are gaining salience in acoustics [52] and ultrasound [53] as a means to combine multiple, usually nonlinear, operations with beamforming. There is also recent interest in beamforming for biomimetic antenna arrays that are based on the direction binaural mechanism of humans or animals [54], [55]. Synthetic apertures across a wide variety of applications, including quantum Rydberg sensing, present unique beamforming challenges [56]. Holographic beamformers [57] are currently investigated as attractive solutions for multibeam steering for future wireless applications. In the following, we illustrate a few major applications in the context of radar and communications.

Joint radar communications

For several decades, sensing and communications systems have exclusively operated in different frequency bands to minimize interference with each other at all times. However, this conservative approach for spectrum access is no longer viable because of the demand for wider bandwidth for the improved performance of both systems. In the last few years, there has been substantial interest in designing joint radar and communications (JRC) [58] to share the spectrum. From a beamformer design perspective, the problem settings of communications and sensing are combined in JRC. Recall the hybrid beamforming for a communications-only problem as explained in (23). The sensing-only beamformer composed of the steering vectors corresponding to, say, K sensing targets is $\mathbf{F}_R \in \mathbb{C}^{N_T \times K}$ [43]. Then, similar to (23), the hybrid beamformer for a sensing-only system is obtained by minimizing the Euclidean distance between $\mathbf{F}_{RF}\mathbf{F}_{BB}$ and $\mathbf{F}_R\mathbf{P}$ as

$$\underset{\mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{P}}{\text{minimize}} \|\mathbf{F}_{RF}\mathbf{F}_{BB} - \mathbf{F}_R\mathbf{P}\|_{\mathcal{F}}$$

$$\text{subject to } \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_{\mathcal{F}} = N_S, \ |[\mathbf{F}_{RF}]_{i,j}| = \frac{1}{\sqrt{N}}, \ \forall i, j, \ \mathbf{P}\mathbf{P}^{\mathsf{H}} = \mathbf{I}_K \tag{25}$$

where the unitary matrix $\mathbf{P} \in \mathbb{C}^{K \times N_S}$ is an auxiliary variable to account for different dimensions of $\mathbf{F}_{RF}\mathbf{F}_{BB}$ and $\mathbf{F}_R$ without causing any distortion in the radar beam pattern. Define $\mathbf{F}_{CR} \in \mathbb{C}^{N_T \times N_S}$ as the unconstrained JRC beamformer

$\mathbf{F}_{CR} = \zeta\mathbf{F}_C + (1 - \zeta)\mathbf{F}_R\mathbf{P}$, where $0 \leq \zeta \leq 1$ provides a tradeoff between radar and communications performance. Then, the JRC hybrid beamformer is obtained by solving the following optimization problem [43]:

$$\underset{\mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{P}}{\text{minimize}} \|\mathbf{F}_{RF}\mathbf{F}_{BB} - \mathbf{F}_{CR}\|_{\mathcal{F}}$$

$$\text{subject to } \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_{\mathcal{F}} = N_S, \ |[\mathbf{F}_{RF}]_{i,j}| = \frac{1}{\sqrt{N}}, \ \forall i, j, \ \mathbf{P}\mathbf{P}^{\mathsf{H}} = \mathbf{I}_K. \tag{26}$$

Radar and communications can be combined in other ways, for example, leveraging the radar information in a different band to reduce the overheads of configuring the beamforming for communication [59].

THz communications

THz-band (0.1–10-THz) wireless systems have ultrawide bandwidth and very narrow beamwidth. The signal processing for these systems must address several unique THz challenges, including severe path loss arising from scattering and molecular absorption. In general, THz communications systems employ ultramassive antenna arrays, which may be variously configured as an array of subarrays or group of subarrays [43] (Figure 4) to achieve even higher beamforming gain than mm-wave systems. The wideband beamforming required at THz uses a single analog beamformer for all subcarriers for a hardware-efficient and computationally inexpensive design. However, this leads to beams generated at the lower and higher subcarriers pointing at different directions, resulting in the beam-squint phenomenon [43]. For comparison's sake, the angular deviation in the beam space due to beam squint is approximately 6° (0.4°) for 0.3 THz with a 30-GHz (60 GHz with a 1-GHz) bandwidth, respectively. One approach to deal with beam squint is to use time-delayer networks, which is classically known as *space–time filtering*. Alternatively, one may design a single analog beamformer while passing the effect of beam squint into the subcarrier digital beamformers.

Consider the problem in (24), where the analog beamformers are subcarrier independent but the mitigation of beam squint implies their SD-ness. Define $\tilde{\mathbf{F}}_{BB}[m]$ as a beam-squint-aware digital beamformer. This is obtained via $\tilde{\mathbf{F}}_{BB}[m] = \mathbf{F}_{RF}^\dagger\bar{\mathbf{F}}_{RF}[m]\mathbf{F}_{BB}[m]$, where $\bar{\mathbf{F}}_{RF}[m]$ is the SD analog beamformer derived from $\mathbf{F}_{RF}$ for $m \in \mathcal{M}$ [43].

Intelligent reflecting surfaces

An intelligent reflecting surface (IRS) is composed of a large number of (usually passive) metamaterial elements, which reflect the incoming signal by introducing a predetermined phase shift [60]. Thus, IRS-assisted beamforming allows the BS to reach distant/blocked users/targets with low power consumption (Figure 4). Here, joint optimization of the beamformers at the BS as well as the phase shifts of IRS elements is necessary. Consider an IRS-assisted scenario, wherein the IRS is equipped with N_{IRS} elements, and the BS has N antennas. The transmitted data symbol $s \in \mathbb{C}$ is received at the user as $y_{IRS} = (\mathbf{h}_{IRS}^{\mathsf{H}}\boldsymbol{\psi}\mathbf{H}_{BS} + \mathbf{h}_D^{\mathsf{H}})\mathbf{f}s + e$, where $\mathbf{h}_{IRS} \in \mathbb{C}^{N_{IRS}}$, $\mathbf{h}_D \in \mathbb{C}^N$, and $\mathbf{H}_{BS} \in \mathbb{C}^{N_{IRS} \times N}$ are the user–IRS, user–BS,

and BS–IRS channels, respectively; the diagonal matrix $\boldsymbol{\psi} = \mathrm{diag}\{[\psi_1, \ldots, \psi_{N_{\mathrm{IRS}}}]\} \in \mathbb{C}^{N_{\mathrm{IRS}} \times N_{\mathrm{IRS}}}$ represents the IRS phase elements; $\mathbf{f} \in \mathbb{C}^N$ is the beamformer vector at the BS; and $e \in \mathbb{C}$ is additive noise. The joint active/passive beamformer design becomes

$$\underset{\boldsymbol{\psi}, \mathbf{f}}{\text{maximize}} \; \left|(\mathbf{h}_{\mathrm{IRS}}^{\mathsf{H}} \boldsymbol{\psi} \mathbf{H}_{\mathrm{BS}} + \mathbf{h}_{\mathrm{D}}^{\mathsf{H}})\mathbf{f}\right|^2$$
$$\text{subject to } \|\mathbf{f}\|_2 \leq \bar{p}, \quad 0 \leq \psi_n \leq 2\pi \qquad (27)$$

where $\bar{p}$ denotes the maximum transmit power, and $n = 1, \ldots, N_{\mathrm{IRS}}$.

Near-field beamforming

Depending on the operating frequency, the wavefront of the transmitted signal appears to have different shapes in accordance with the observation distance. The wavefront is a plane wave in the far-field region. In the near field (Figure 4), where the transmission range is shorter than the Fraunhofer distance, i.e.,

$$R_{\mathrm{NF}} = \frac{2A^2 f_c}{c_0}$$

with A being the array aperture, the wavefront takes a spherical form. As a result, unlike the far field, the near-field beam pattern is range dependent. For example, the array response vector for uniform linear array (ULA) is a function of both direction θ and range r as

$$\mathbf{a}(\theta, r) = \frac{1}{\sqrt{N}} \left[e^{-\mathrm{j}\frac{2\pi}{\lambda} r^{(1)}}, \ldots, e^{-\mathrm{j}\frac{2\pi}{\lambda} r^{(N)}} \right]^{\mathsf{T}}$$

where $r^{(n)} = [r^2 + ((n-1)d)^2 - 2(n-1)dr\sin\theta]^{1/2} \approx r - (n-1)$ $d\sin\theta$, $(n = 1, \ldots, N)$ is a range-dependent parameter corresponding to the receiver and the nth transmit antenna. Hence, the beamformer design needs to account for this spherical model.

Summary

The many beamforming algorithms, their possible variants, and their relative advantages provide a Swiss-knife approach to choosing the most appropriate technique for a specific application. We presented an overview of those algorithms that had a considerable impact on signal processing and system design during the last 25 years. We focused on radar and communications applications while also mentioning in passing the developments in beamforming for ultrasound, acoustics, synthetic apertures, and optics.

A typical use case of convex beamforming is to allow robustness against various sources of uncertainties, such as a small number of snapshots, mismatched SoI direction, and mismatched steering vectors. In nonconvex beamforming, each of the problem settings imposes different constraints on, e.g., PSD-ness (general-rank beamforming), the probability distribution (chance-constrained robust beamforming), constant-modulus (hybrid beamforming), and received SNR (multicast beamforming).

Each learning algorithm offers specific advantages of its own. The most common SL (UL and RL) admits labeled (unlabeled) datasets. Furthermore, the inherent reward/punishment mechanism in RL to optimize the learning model for a predefined cost function yields better performance than UL. FL is particularly helpful for multiuser scenarios, whereas CL is preferred if the dataset is small compared to the size of the learning model. When data are updated over time, then OL is beneficial. Note that SL, UL, and RL may also be combined with FL, CL, and OL depending on the problem and data; examples abound, such as federated RL, online RL, online CL, centralized RL, and so on.
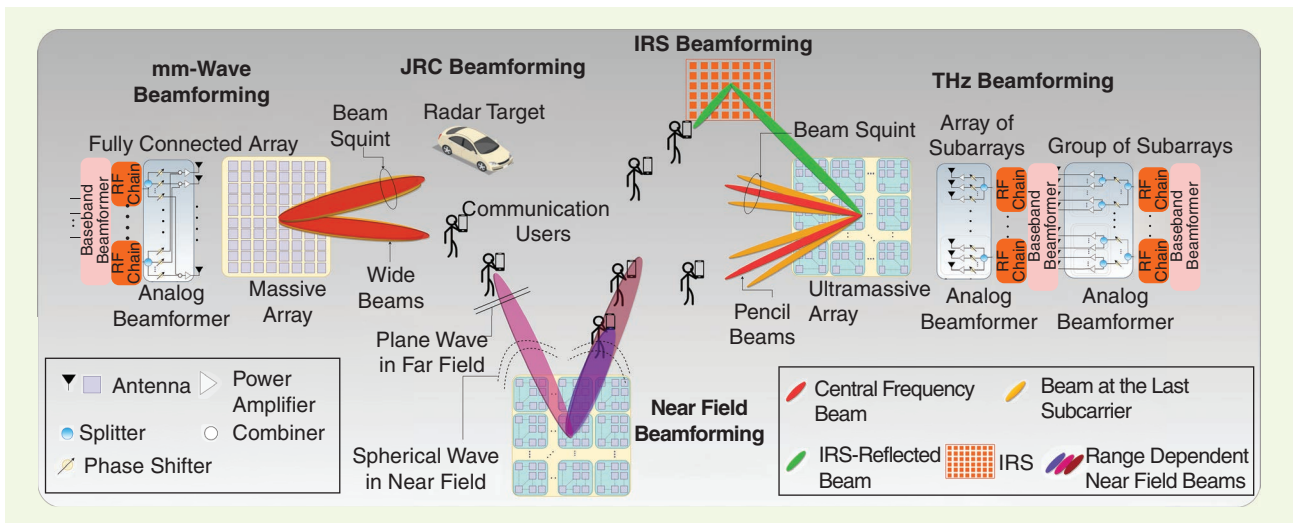
FIGURE 4. A summary of beamforming in emerging applications. In mm-wave wideband beamforming, the generated beams are squinted while pointing to the same direction, but, at THz, these beams are squinted in considerably different directions. In a JRC scenario, a joint optimization of the beam pattern for both communications users and radar targets should be considered. For IRS-assisted wireless systems, the beamformer weights at the transmitter and the phase shifts of the IRS elements are jointly designed. When the users are in the near-field region of the transmitter, range-dependent beamforming is considered for spatial multiplexing.

Acknowledgment

Authors

Ahmet M. Elbir (ahmetmelbir@ieee.org) received his Ph.D. degree from the Middle East Technical University (METU), Turkey, in 2016 in electrical engineering. He is a research fellow at the University of Luxembourg, L-1855 Luxembourg City, Luxembourg, and senior researcher at Duzce University, Duzce, 81620, Turkey. He serves as an associate editor for *IEEE Access* and a lead guest editor for *IEEE Journal of Selected Topics in Signal Processing* and *IEEE Wireless Communications*. He is the recipient of the 2016 METU Best Ph.D. Thesis Award for his doctoral studies and the IET Radar, Sonar, and Navigation Best Paper Award in 2022. His research interests include array signal processing for radar and communications as well as deep learning for multiantenna systems. He is a Senior Member of IEEE.

Kumar Vijay Mishra (kvm@ieee.org) received his Ph.D. degree in electrical and computer engineering from the University of Iowa while working on the NASA Global Precipitation Measurement Mission ground validation radars. He is a senior fellow at the U.S. DEVCOM Army Research Laboratory, Adelphi MD 20783 USA, and technical advisor to start-ups Hertzwell, Singapore, and Aura Intelligent Systems, Boston, MA USA. He is the recipient of the U.S. National Academies Harry Diamond Distinguished Fellowship and has won many best paper awards. His research interests include radar, remote sensing, signal processing, and electromagnetics. He is a Senior Member of IEEE.

Sergiy A. Vorobyov (svor@ieee.org) received his Ph.D. degree in systems and control from the National University of Radio Electronics, Kharkiv, Ukraine. He is a professor with the Department of Information and Communications Engineering, Aalto University, 02150 Espoo, Finland. He was the recipient of the 2004 IEEE Signal Processing Society Best Paper Award, 2007 Alberta Ingenuity New Faculty Award, 2011 Carl Zeiss Award, 2012 NSERC Discovery Accelerator Award, and other awards. He is currently serving as the general cochair for EUSIPCO 2023, Helsinki, Finland. His research interests include optimization and multilinear algebra methods in signal processing and data analysis; statistical and array signal processing; sparse signal processing; estimation; detection and learning theory and methods; and multiantenna, large-scale, and cognitive systems. He is a Fellow of IEEE.

Robert W. Heath Jr. (rwheathjr@ncsu.edu) received his Ph.D. degree from Stanford University in electrical engineering. He is the Lampe Distinguished Professor at North Carolina State University, Raleigh, NC 27695 USA, and is president and CEO of MIMO Wireless Inc. He has authored or coauthored several books, including *Introduction to Wireless Digital Communication* (Prentice Hall, 2017) and *Foundations of MIMO Communication* (Cambridge University Press, 2018). He is the recipient or corecipient of several awards, including the 2019 IEEE Kiyo Tomiyasu Award, 2020 IEEE Signal Processing Society Donald G. Fink Overview Paper Award, 2020 North Carolina State University Innovator of the Year Award, 2021 IEEE Vehicular Technology Society James Evans Avant Garde Award, and 2022 IEEE Vehicular Technology Society Best Vehicular Electronics Paper Award. He was editor-in-chief of *IEEE Signal Processing Magazine* from 2018 to 2020. He is a fellow of the National Academy of Inventors and a Fellow of IEEE.

References

[1] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988, doi: 10.1109/53.665.

[2] R. Simons, "Guglielmo Marconi and early systems of wireless communication," *Gec Rev.*, vol. 11, no. 1, pp. 37–55, Jan. 1996.

[3] T. K. Sarkar, R. Mailloux, A. A. Oliner, M. Salazar-Palma, and D. L. Sengupta, *History of Wireless*. Hoboken, NJ, USA: Wiley, 2006.

[4] F. Bartlett, "A dual diversity preselector," *QST*, vol. 25, pp. 37–39, Apr. 1941.

[5] J. C. Chen and K. Yao, "Beamforming," in *Distributed Sensor Networks: Image and Sensor Signal Processing*, vol. 2, S. S. Iyengar and R. R. Brooks, Eds. Boca Raton, FL, USA: CRC Press, 2016, pp. 335–371.

[6] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969, doi: 10.1109/PROC.1969.7278.

[7] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive antenna systems," *Proc. IEEE*, vol. 55, no. 12, pp. 2143–2159, Dec. 1967, doi: 10.1109/PROC.1967.6092.

[8] S. A. Vorobyov, "Adaptive and robust beamforming," in *Array and Statistical Signal Processing* (Academic Press Library in Signal Processing), vol. 3, A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. New York, NY, USA: Academic Press, 2014, pp. 503–552.

[9] H. Cox, "Resolving power and sensitivity to mismatch of optimum array processors," *J. Acoust. Soc. Amer.*, vol. 54, no. 3, 2005, Art. no. 771, doi: 10.1121/1.1913659.

[10] N. Jablon, "Adaptive beamforming with the generalized sidelobe canceller in the presence of array imperfections," *IEEE Trans. Antennas Propag.*, vol. 34, no. 8, pp. 996–1012, Aug. 1986, doi: 10.1109/TAP.1986.1143936.

[11] A. B. Gershman, V. I. Turchin, and V. A. Zverev, "Experimental results of localization of moving underwater signal by adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 43, no. 10, pp. 2249–2257, Oct. 1995, doi: 10.1109/78.469863.

[12] D. Astely and B. Ottersten, "The effects of local scattering on direction of arrival estimation with MUSIC," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3220–3234, Dec. 1999, doi: 10.1109/78.806068.

[13] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003, doi: 10.1109/TSP.2002.806865.

[14] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1365–1376, Oct. 1987, doi: 10.1109/TASSP.1987.1165054.

[15] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003, doi: 10.1109/TSP.2003.812831.

[16] D. D. Feldman and L. J. Griffiths, "A projection approach for robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 42, no. 4, pp. 867–876, Apr. 1994, doi: 10.1109/78.285650.

[17] S. Shahbazpanahi, A. B. Gershman, Z.-Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2257–2269, Sep. 2003, doi: 10.1109/TSP.2003.815395.

[18] A. Khabbazibasmenj and S. A. Vorobyov, "Robust adaptive beamforming for general-rank signal model with positive semi-definite constraint via POTDC," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6103–6117, Dec. 2013, doi: 10.1109/TSP.2013.2281301.

[19] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010, doi: 10.1109/MSP.2010.936015.

[20] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005, doi: 10.1109/TSP.2005.845436.

[21] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust Capon beamformer," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2407–2423, Sep. 2004, doi: 10.1109/TSP.2004.831998.

[22] Y. Huang, M. Zhou, and S. Vorobyov, "New designs on MVDR robust adaptive beamforming based on optimal steering vector estimation," *IEEE Trans. Signal Process.*, vol. 67, no. 14, pp. 3624–3638, Jul. 2019, doi: 10.1109/TSP.2019.2918997.

[23] A. Hassanien, S. Vorobyov, and K. Wong, "Robust adaptive beamforming using sequential quadratic programming: An iterative solution to the mismatch problem," *IEEE Signal Process. Lett.*, vol. 15, pp. 733–736, Nov. 2008, doi: 10.1109/LSP.2008.2001115.

[24] A. Khabbazibasmenj, A. Hassanien, and S. Vorobyov, "Robust adaptive beamforming based on steering vector estimation with as little as possible prior information," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2974–2987, Jun. 2012, doi: 10.1109/TSP.2012.2189389.

[25] S. Vorobyov, H. Chen, and A. Gershman, "On the relationship between robust minimum variance beamformers with probabilistic and worst-case distortionless response constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5719–5724, Nov. 2008, doi: 10.1109/TSP.2008.929866.

[26] Y. Huang, W. Yang, and S. A. Vorobyov, "Robust adaptive beamforming maximizing the worst-case SINR over distributional uncertainty sets for random INC matrix and signal steering vector," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2022, pp. 4918–4922, doi: 10.1109/ICASSP43922.2022.9746616.

[27] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jul. 2006, doi: 10.1109/TSP.2006.872578.

[28] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014, doi: 10.1109/TWC.2014.011714.130846.

[29] Y. Savas, E. Noorani, A. Koppel, J. Baras, U. Topcu, and B. M. Sadler, "Collaborative one-shot beamforming under localization errors: A discrete optimization approach," *Signal Process.*, vol. 200, Nov. 2022, Art. no. 108647, doi: 10.1016/j.sigpro.2022.108647.

[30] Ö. T. Demir and T. E. Tuncer, "Alternating maximization algorithm for the broadcast beamforming," in *Proc. 22nd Eur. Signal Process. Conf. (EUSIPCO)*, 2014, pp. 1915–1919.

[31] R. W. Heath Jr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016, doi: 10.1109/JSTSP.2016.2523924.

[32] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016, doi: 10.1109/JSTSP.2016.2523903.

[33] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016, doi: 10.1109/JSTSP.2016.2520912.

[34] A. Alkhateeb and R. W. Heath Jr., "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016, doi: 10.1109/TCOMM.2016.2549517.

[35] S. Mohammadzadeh, V. H. Nascimento, R. C. de Lamare, and N. Hajarolasvadi, "Robust beamforming based on complex-valued convolutional neural networks for sensor arrays," *IEEE Signal Process. Lett.*, vol. 29, pp. 2108–2112, Oct. 2022, doi: 10.1109/LSP.2022.3212637.

[36] A. M. Elbir, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, "A family of deep learning architectures for channel estimation and hybrid beamforming in multi-carrier mm-Wave massive MIMO," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 2, pp. 642–656, Jun. 2022, doi: 10.1109/TCCN.2021.3132609.

[37] A. M. Elbir and K. V. Mishra, "Joint antenna selection and hybrid beamformer design using unquantized and quantized deep learning networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1677–1688, Mar. 2020, doi: 10.1109/TWC.2019.2956146.

[38] P. Dong, H. Zhang, and G. Y. Li, "Framework on deep learning-based joint hybrid processing for mmWave massive MIMO systems," *IEEE Access*, vol. 8, pp. 106,023–106,035, Jun. 2020, doi: 10.1109/ACCESS.2020.3000601.

[39] A. Beck and Y. C. Eldar, "Doubly constrained robust capon beamformer with ellipsoidal uncertainty sets," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 753–758, Jan. 2007, doi: 10.1109/TSP.2006.885729.

[40] X. Jiang, W.-J. Zeng, A. Yasotharan, H. C. So, and T. Kirubarajan, "Minimum dispersion beamforming for non-gaussian signals," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1879–1893, Apr. 2014, doi: 10.1109/TSP.2014.2305639.

[41] A. Parayil, A. S. Bedi, and A. Koppel, "Joint position and beamforming control via alternating nonlinear least-squares with a hierarchical gamma prior," in *Proc. Amer. Control Conf. (ACC)*, 2021, pp. 3513–3518, doi: 10.23919/ACC50511.2021.9482851.

[42] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017, doi: 10.1109/TSP.2017.2706179.

[43] A. M. Elbir, K. V. Mishra, and S. Chatzinotas, "Terahertz-band joint ultra-massive MIMO radar-communications: Model-based and model-free hybrid beamforming," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1468–1483, Nov. 2021, doi: 10.1109/JSTSP.2021.3117410.

[44] M. A. B. Abbasi, V. F. Fusco, H. Tataria, and M. Matthaiou, "Constant-ϵ_r lens beamformer for low-complexity millimeter-wave hybrid MIMO," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 2894–2903, Jul. 2019, doi: 10.1109/TMTT.2019.2903790.

[45] E.-A. Fazal, C. C. Cavalcante, F. Antreich, A. L. F. De Almeida, and J. A. Nossek, "Efficient hybrid A/D beamforming for millimeter-wave systems using butler matrices," *IEEE Trans. Wireless Commun.*, vol. 22, no. 2, pp. 1001–1013, Feb. 2023, doi: 10.1109/TWC.2022.3200298.

[46] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath Jr., "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014, doi: 10.1109/MCOM.2014.6979963.

[47] A. M. Elbir, "CNN-based precoder and combiner design in mmWave MIMO systems," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1240–1243, Jul. 2019, doi: 10.1109/LCOMM.2019.2915977.

[48] S. Wager, A. Khare, M. Wu, K. Kumatani, and S. Sundaram, "Fully learnable front-end for multi-channel acoustic modeling using semi-supervised learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2020, pp. 6864–6868, doi: 10.1109/ICASSP40776.2020.9053367.

[49] Q. Wang, K. Feng, X. Li, and S. Jin, "PrecoderNet: Hybrid beamforming for millimeter wave systems with deep reinforcement learning," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1677–1681, Oct. 2020, doi: 10.1109/LWC.2020.3001121.

[50] A. M. Elbir and S. Coleri, "Federated learning for hybrid beamforming in mm-Wave massive MIMO," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2795–2799, Dec. 2020, doi: 10.1109/LCOMM.2020.3019312.

[51] S. Shi, Y. Cai, Q. Hu, B. Champagne, and L. Hanzo, "Deep-unfolding neural-network aided hybrid beamforming based on symbol-error probability minimization," *IEEE Trans. Veh. Technol.*, vol. 72, no. 1, pp. 529–545, Jan. 2023, doi: 10.1109/TVT.2022.3201961.

[52] T. Nakatani and K. Kinoshita, "A unified convolutional beamformer for simultaneous denoising and dereverberation," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 903–907, Jun. 2019, doi: 10.1109/LSP.2019.2911179.

[53] B. Heriard-Dubreuil, A. Besson, F. Wintzenrieth, J.-P. Thiran, and C. Cohen-Bacrie, "Sparse convolutional plane-wave compounding for ultrasound imaging," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, 2020, pp. 1–4, doi: 10.1109/IUS46767.2020.9251493.

[54] A. R. Masoumi, Y. Yusuf, and N. Behdad, "Biomimetic antenna arrays based on the directional hearing mechanism of the parasitoid fly Ormia Ochracea," *IEEE Trans. Antennas Propag.*, vol. 61, no. 5, pp. 2500–2510, May 2013, doi: 10.1109/TAP.2013.2245091.

[55] A. R. Masoumi and N. Behdad, "An improved architecture for two-element biomimetic antenna arrays," *IEEE Trans. Antennas Propag.*, vol. 61, no. 12, pp. 6224–6228, Dec. 2013, doi: 10.1109/TAP.2013.2281352.

[56] P. Vouras et al., "An overview of advances in signal processing techniques for classical and quantum wideband synthetic apertures," *IEEE J. Sel. Topics Signal Process.*, early access, Mar. 2023, doi: 10.1109/JSTSP.2023.3262443.

[57] R. Deng, B. Di, H. Zhang, Y. Tan, and L. Song, "Reconfigurable holographic surface: Holographic beamforming for metasurface-aided wireless communications," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6255–6259, Jun. 2021, doi: 10.1109/TVT.2021.3079465.

[58] K. V. Mishra, M. R. Bhavani Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, "Toward millimeter wave joint radar-communications: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 100–114, Sep. 2019, doi: 10.1109/MSP.2019.2913173.

[59] A. Ali, N. Gonzalez-Prelcic, and A. Ghosh, "Passive radar at the roadside unit to configure millimeter wave vehicle-to-infrastructure links," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14,903–14,917, Dec. 2020, doi: 10.1109/TVT.2020.3027636.

[60] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019, doi: 10.1109/TWC.2019.2936025.

SP