Limitations and Implementation Strategies of Interstage Matching in a 6-W, 28–38-GHz GaN Power Amplifier MMIC

Philipp Neininger, *Student Member, IEEE*, Laurenz John, Fabian Thome,
Christian Friesicke, *Member, IEEE*, Peter Brückner, Rüdiger Quay, *Senior Member, IEEE*,
and Thomas Zwick, *Fellow, IEEE*

Abstract— In this article, we summarize the theoretical matching boundaries and show the limitations they implicate for real-world amplifier design. Starting with a common schematic prototype, we investigate the question of how to realize its electrical response in a densely routed, massively parallelized layout. To that end, we develop a comprehensive study on the application of space-mapping techniques toward the design of high-power amplifiers (HPAs). We derive three reference design procedures and compare their performance in terms of convergence, speed, and practicality when laying out a densely routed HPA interstage matching network. Subsequently, we demonstrate the usefulness of the study by designing the networks of a compact three-stage eight-way wideband HPA in the Ka-band. The processed monolithic microwave integrated circuit features a 1-dB large-signal bandwidth of more than 11 GHz (a fractional bandwidth of 32.8%) and thus covers most of the Ka-band with an output power exceeding 6 W in 3 dB of gain compression. This demonstrates the highest combination of power and bandwidth to date using a reactively matched topology in the Ka-band.

Index Terms— Gallium nitride (GaN), high-electron-mobility transistors (HEMTs), Ka-band, millimeter wave (mmW), monolithic microwave integrated circuits (MMICs), power amplifiers (PAs), space mapping, wideband matching.

I. INTRODUCTION

IN RECENT years, as solid-state technologies continue to increase their output power density and efficiency at millimeter-wave frequencies, solid-state amplifiers can be employed in applications that could previously only be realized using traveling-wave tube amplifiers (TWTAs). These applications include intersatellite communications, medical imaging, radar, and mobile communications. In the future, solid-state technology could allow engineers to overcome common drawbacks of traditional TWTAs [1], for example, their warm-up time, limited service life, and their rather narrow usable bandwidth. These issues—especially the last point—can be improved upon considerably using III–V semiconductors, such as gallium arsenide or gallium nitride (GaN). In the Ka-band (26.5–40 GHz), research toward higher power GaN high-power amplifiers (HPAs) recently produced publications demonstrating up to 40 W of output power [2], [3] on a single monolithic microwave integrated circuit (MMIC).

High-power and high-efficiency designs typically utilize a reactive matching network to obtain a large-signal impedance match of the high-electron-mobility transistors (HEMTs) employed in the circuit. These types of MMICs generally achieve up to about 20% of relative bandwidth (RBW) [4], [5]. On the other hand, using traveling-wave topologies, fractional bandwidths exceeding 50% have been shown [6], [7]. However, MMICs employing traveling-wave topologies may exhibit certain disadvantages, among them an uneven heat distribution [6] and comparatively low-efficiency figures [8]. These are compelling reasons to strive for a reactive-type MMIC that reaches a relatively large bandwidth and still maintains reasonable power and efficiency figures.

When designing a matching network in a large and densely routed HPA MMIC, extensive coupling of neighboring structures as well as other distributed effects can dominate the RF-response of the network. Since these effects are not accounted for in circuit models, it can be challenging to transfer a prototype (schematic) matching network to a layout such that the electrical responses are sufficiently similar, especially over a large bandwidth. Once we approach the fundamental matching boundary given by the Bode–Fano limit [9], [10], the accuracy to which the prototype network must be implemented is even more critical. This is because in a design that is close to the theoretical limit, additional parasitics will deteriorate the obtainable match inside the desired frequency range. In contrast to this, in a design that is considerably below the theoretical limit, additional parasitic elements can be absorbed without affecting the theoretically obtainable match. Consequently, dedicated techniques are useful to systematically tweak a layout to approach the originally desired char-

acteristic. Space mapping [11]–[15] is one such technique that has gained significant use in the design of filters [16]. While its usage in MMIC design has been reported before [17], this is the first article showing a detailed study on the properties and usefulness of this concept toward high-power MMIC design. We concentrate on the practical implementation of wideband interstage matching structures. To that end, we derive, implement, and compare two different algorithms: one of them more complex and close to the full formulation of space mapping [11] and a second one that features a reduced number of steps and is simpler to implement in an EDA program such as Keysight ADS.

This article is organized as follows. In Section II, we introduce Fraunhofer IAF's GaN10 process and describe the properties of HEMT devices manufactured with it. Section III describes the theoretical boundaries of matching networks and examines the challenges arising from interstage matching networks (ISMN), where a complex impedance is matched to another complex impedance. A common topology of choice for an ISMN is described in Section IV. Section V contains a discussion of implementation approaches. The space-mapping technique is described, followed by the derivation of two concrete algorithms, a reference implementation, and a comparison in terms of performance and usability. Sections VI and VII give details on a wideband HPA MMIC design which was implemented using the study above. A comparison of the processed MMIC to the state of the art is given in Section VIII. Finally, the conclusion and outlook can be found in Section IX.

II. TECHNOLOGY AND DEVICE

Fraunhofer IAF's GaN10 technology features AlGaN/GaN-HEMTs with 100-nm T-gates. As a substrate, 100-mm semi-insulating silicon carbide with a thickness of $75\,\mu\text{m}$ is used. Passive circuitry is realized using an evaporated first and a galvanic second metal layer for increased current handling capability. Metal–insulator–metal (MIM) capacitors and nickel–chromium (NiCr) thin-film resistors are available as well as a full backside process with through-substrate vias. The transition frequency f_T and maximum oscillation frequency f_{max} of this process are in the range of 100 and 300 GHz, respectively [18]. This enables MMICs operating at W-band and beyond [19]. With a typical output power density of $2\,\text{W/mm}$ at $V_{DS} = 15\,\text{V}$, high-power designs exhibiting state-of-the-art efficiency can be realized.

For this research, we analyzed the perspective of broadband parallel power combining in the lower mmW frequency band. A previous design has shown promising results in the same frequency band [20] and is improved upon in this work. As shown in [20], the optimum HEMT periphery for the frequency band in question was found to be an eight finger device with a unit gate width of UGW = $60\,\mu\text{m}$ for a total gate width of TGW = $0.48\,\text{mm}$. Therefore, they are used for this work as well.

A major challenge in the design of broadband MMICs is to find a matching network that compensates for (in fact, absorbs) the parasitic capacitances, these devices exhibit at high frequencies. To that end, the HEMT gate can typically
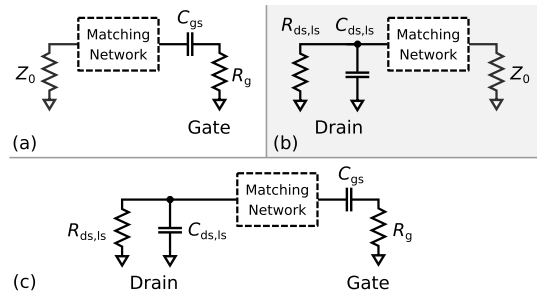

Fig. 1. Matching of simplified gate and drain equivalent circuits. (a) Equivalent circuit of the gate side (input matching). (b) Equivalent circuit of the transistor drain (output matching). (c) Concept of interstage matching, where both the input and the output capacitances have to be absorbed by the matching network.
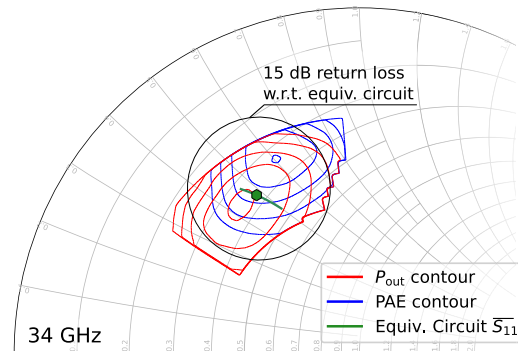

Fig. 2. Device-level LP measurement results at 34 GHz. The red and blue shapes indicate contours of constant P_{out} and PAE, respectively. The equivalent circuit's $\overline{S_{11}}$ between 28 and 38 GHz is shown in green (● marks its value at 34 GHz). The target match of $\Gamma_m \leq -15$ dB is depicted as a black circle.

be modeled as a series (L)CR element, while its drain behaves like a parallel RC circuit [21]. Both of these equivalent circuits are labeled accordingly in Fig. 1. To derive appropriate values for the large-signal equivalent circuits, load–pull (LP) measurements [22] are a valuable basis. As an example, Fig. 2 is generated from an LP measurement at $f_0 = 34\,\text{GHz}$. It shows the contours of constant P_{out} and PAE in 2-dB gain compression. Each contour represents a step of 50 mW and 2% in P_{out} and PAE, respectively. By selecting appropriate values for $R_{ds,ls}$ and $C_{ds,ls}$, we can synthesize matching networks that result in maximum PAE or maximum P_{out} or choose a tradeoff goal. For a general-purpose PA, a tradeoff impedance in between the two maximum points is most sensible. In this way, we can ensure to obtain as robust a design as possible and allow for measurement uncertainty and process variations, which could shift the optimum impedance point slightly.

Using LP measurements at 30 and 34 GHz, we deduce the following values for the optimum device operating at $V_{ds} = 15\,\text{V}$: the gate capacitance equals $C_{gs} = 443\,\text{fF}$ and the gate resistance is $R_g = 3.32\,\Omega$. For the HEMT drain, we find $C_{ds,ls} = 122\,\text{fF}$ and the drain loadline resistance $R_{ds,ls} = 55.5\,\Omega$. $\overline{S_{11}}$ of the resulting drain equivalent circuit is shown as a green trace in Fig. 2 (28–38 GHz). The goal in the following will be to design an ISMN that achieves 85% of the maximum P_{out} and PAE in the frequency band of interest.

This equals an impedance range as depicted by the black circle in Fig. 2 around the tradeoff impedance, which represents a maximum reflection coefficient of $\Gamma_m \leq -15$ dB with respect to the tradeoff impedance.

III. THEORETICAL MATCHING BOUNDARY

Given the passive equivalent circuit of an HEMT device, it turns out that there is an upper limit as to how broadband it can be matched in principle. The Bode–Fano criterion defines this limit, which is the maximum matching bandwidth that can be achieved for a given combination of capacitance and termination resistance [23]. In this context, we define *matching bandwidth* as the frequency range in which the reflection coefficient Γ is below a certain threshold Γ_m.

For an RC-parallel element as shown in Fig. 1(b), the reflection coefficient Γ is limited by the inequality [9]

$$\int_0^\infty \ln \frac{1}{|\Gamma(\omega)|} \, d\omega < \frac{\pi}{RC} = \frac{\pi}{\tau} \tag{1}$$

where τ is the RC time constant [24]. Assuming a constant reflection over the band of interest and total reflection out of band, we can simplify (1) to

$$(\omega_1 - \omega_2) \ln \Gamma_m < \frac{\pi}{\tau} \tag{2}$$

which gives us an upper boundary of the maximum achievable reflection coefficient. In general, the RC time constant τ is inversely proportional to the total area under the reflection coefficient curve (i.e., S_{xx} in dB).

Furthermore, the Bode–Fano limit for the HEMT input equivalent circuit, shown in Fig. 1 on the left, is given by

$$\int_0^\infty \frac{1}{\omega^2} \ln \frac{1}{|\Gamma(\omega)|} \, d\omega < \pi \tau. \tag{3}$$

As before, we can simplify this equation assuming a constant in-band reflection coefficient

$$\ln \Gamma_m \cdot \frac{\omega_1 - \omega_2}{\omega_1 \omega_2} < \pi \tau. \tag{4}$$

Introducing fractional bandwidth, RBW, as

$$\text{RBW} = 2 \frac{\omega_2 - \omega_1}{\omega_1 + \omega_2} \tag{5}$$

and the center frequency ω_c as

$$\omega_c = \sqrt{\omega_1 \omega_2} \tag{6}$$

we can rearrange (4) to

$$\frac{1}{\omega_c} \cdot \frac{4 \ln \Gamma_m \cdot \text{RBW}}{\text{RBW}^2 - 4} < \pi \tau. \tag{7}$$

In contrast to the HEMT drain equivalent circuit (2), the result of (7) illustrates an improved match for higher center frequencies ω_c. In other words, for the gate, the achievable RBW or alternatively Γ_m improves with frequency, while for the drain, it decreases.

Due to their higher operating voltage, high-voltage technologies such as GaN feature HEMTs with high $R_{ds,ls}$ and are as such intrinsically more limited in terms of their Bode–Fano matching bandwidth compared to GaAs pHEMT or even silicon devices [25]. An example of the Bode–Fano limit of
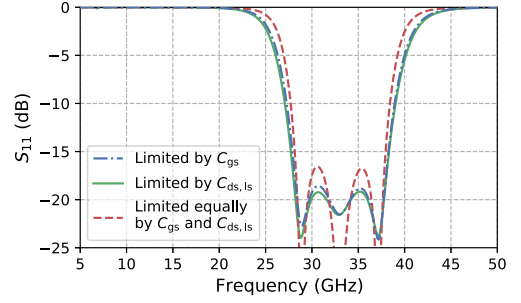

Fig. 3. S_{11} of a lumped bandpass matching network that matches a parallel RC circuit to a series RC circuit [as shown in Fig. 1(c)]. Three different cases are shown where we varied the values of C_{gs} and $C_{ds,ls}$ such that the matching is Bode–Fano limited either on the drain side (—), the gate side (− ·), or both (− −). For each of the cases, an optimization was performed to minimize the in-band reflection coefficient.

a low-voltage technology can be calculated using the values supplied in [26], where at $V_{ds} = 1$ V, a two-finger device featuring an UGW of 45 μm exhibits $R_{ds,ls} = 12.5 \, \Omega$ and $C_{ds,ls} = 120$ fF. With the values for GaN10 at $V_{ds} = 15$ V given in Section II, we can compare the Bode–Fano limit for the technologies using (2). The comparison shows that the theoretical limit for GaN10 is roughly 22% of the aforementioned low-voltage technology. In other words, the drain of the low-voltage technology can be matched over 4.5 times the bandwidth with the same reflection coefficient as that of the high-voltage technology. Note that the transistor size is irrelevant to this boundary, as the HEMT's RC time constant τ stays approximately constant over TGW.

As it turns out, in terms of design complexity, the most critical matching network of a GaN high-power amplifier (HPA) with multiple stages is the ISMN. Its practical realization and the tradeoffs required will be described in detail in the following. From a conceptual point of view, we can use the previously introduced HEMT equivalent circuits to represent this problem, as shown in Fig. 1. Fig. 1(a) and (b) shows the matching of the input and output, respectively, whereas Fig. 1(c) shows the case of an interstage network. This arrangement, a frequency-dependent generator impedance to be matched to a frequency-dependent load impedance, is called a *double-matching problem*. Analytical and numerical topology synthesis procedures have been studied extensively in the past [27]–[29].

The fundamental matching limit of the formulation by Bode [9] refers to a purely resistive generator impedance that is matched to an RC load. Fano [10] extended this formulation to a load of arbitrary but fixed impedance, again with a purely resistive generator impedance. However, considering the ISMN's matching bandwidth limit, there is a complex source impedance that is matched to a complex load impedance [see Fig. 1(c)]. Effectively, both the Bode–Fano limit for the drain side [parallel RC, (1)] and for the gate side [series RC, (3)] have to be satisfied.

To investigate a possible penalty in the maximum matching bandwidth when both termination impedances are complex, we compared three different cases in which we varied the

TABLE I
COMPARISON OF GATE- AND DRAIN-LIMITED MATCHING

Matching limited	C_{gs} (fF)	$C_{\mathrm{ds,ls}}$ (fF)	$f_{\mathrm{BF,gs}}$ (%)	$f_{\mathrm{BF,ds}}$ (%)
Only By C_{gs}	26.5	0	95.6	-
Only By $C_{\mathrm{ds,ls}}$	∞	292	-	98.6
By both C_{gs} and $C_{\mathrm{ds,ls}}$	26.5	292	93.7	93.4

Note: $f_{\mathrm{BF,gs}}$ and $f_{\mathrm{BF,ds}}$ are the fraction of the theoretical Bode-Fano limit for gate- and drain that is reached by the network shown in Fig. 3, respectively.

values of the capacitors from Fig. 1(c). The first two cases are those that occur in an input and an output matching network: they are Bode–Fano limited on the gate (C_{gs}) or drain ($C_{\mathrm{ds,ls}}$) side, respectively. For the third case, C_{gs} and $C_{\mathrm{ds,ls}}$ are chosen such that they both equally limit the matching bandwidth. A third-order lumped bandpass was used as the matching network and numerically optimized to exhibit the lowest possible S_{11} between 28 and 38 GHz. As a result, the curves shown in Fig. 3 were obtained. By integrating the area under the curves from Fig. 3 as specified in (1) and (3), we can quantify how close the networks come to the theoretical optimum case, which is given by the right-hand side of (1) and (3). We define the ratio of these experimental values to the Bode–Fano limit as $f_{\mathrm{BF,ds}}$ and $f_{\mathrm{BF,gs}}$

$$f_{\mathrm{BF,gs}} = \frac{1}{\pi \tau} \cdot \int_0^\infty \frac{1}{\omega^2} \ln \frac{1}{|\Gamma(\omega)|} \, d\omega \tag{8}$$

and

$$f_{\mathrm{BF,ds}} = \frac{\tau}{\pi} \cdot \int_0^\infty \ln \frac{1}{|\Gamma(\omega)|} \, d\omega. \tag{9}$$

The resulting values for $f_{\mathrm{BF,ds}}$ and $f_{\mathrm{BF,gs}}$ as well as the capacitance values for each case are listed in Table I. The integration has been performed for frequencies in the range of 0–70 GHz. In the first case, which is limited by C_{gs}, the optimized network reaches $f_{\mathrm{BF,gs}} = 95.6\%$ of the Bode–Fano limit. For the second case (limited by C_{ds}), we obtain a value of $f_{\mathrm{BF,ds}} = 98.6\%$. In the third case, limited by both C_{gs} and C_{ds}, the network reaches 93.7% and 93.4% of the Bode–Fano limits for series and parallel RC circuits, respectively. These results suggest that the used bandpass network can absorb both the series and the parallel capacitance equally well. In addition, the presence of a second constraint (i.e., both terminations are complex), as shown in case 3, has only a marginal impact on the values of $f_{\mathrm{BF,ds}}$ and $f_{\mathrm{BF,gs}}$. For a practical design, we can state that depending on the device's equivalent circuit values, either the gate or drain can limit the obtainable matching performance, depending on which of them is closer to the Bode–Fano limit.

For the following investigation, we assume a design goal of $\Gamma < -15$ dB to the optimum large-signal load targets in the frequency band between 28 and 38 GHz. As a numerical example assuming a network of infinite order (brick-wall response), we obtain from (2) for the drain side

$$2\pi \, 10\,\mathrm{GHz} \cdot \ln(-15\,\mathrm{dB}) = 1.1 \times 10^{11} < 4.6 \times 10^{11}. \tag{10}$$
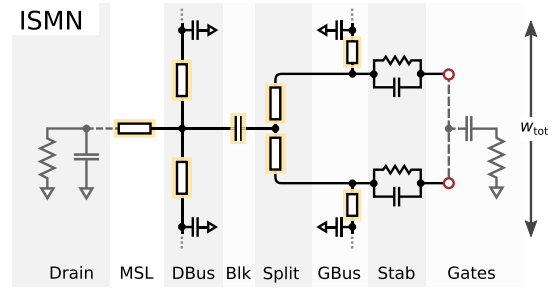

Fig. 4. Illustration of the simulation model for an ISMN base cell. It includes the first stage FET parasitics (*Drain*), the drain bus (*DBus*) with bias insertion points, a dc blocking capacitor *Blk*, a power splitter *Split*, the gate biasing bus (*GBus*), a stabilization network *Stab*, and finally the gate parasitics of the next stage's FET(s).

For the gate side, we find from (4)

$$2.58 \times 10^{-12} < 4.62 \times 10^{-12}. \tag{11}$$

These results indicate that we are limited by the gate side (below the limit by a factor of 1.8), whereas the drain-side matching limit is significantly above the design parameters required (factor of 4.1). Note that for this analysis, the networks are assumed to be loss-less, which, especially for mmW frequencies, will not be the case. The losses introduced either on purpose (e.g., stabilization) or due to parasitics will change the network's return loss (RL).

IV. INTERSTAGE MATCHING TOPOLOGY

In an HPA using parallel power combining, certain components are needed as a consequence of the physical and electrical constraints in place—irrespective of the matching function, which the network has to provide as well.

Fig. 4 shows those components for the case of an ISMN. As seen from the HEMT drain, a bus interconnecting the parallel stages' drains is needed for bias supply. A series capacitor separates the dc path for the gate and drain supply voltages. An n-way power splitter divides the energy toward the next stage and provides the lateral distance required for heat distribution. Furthermore, a gate supply bus is needed to set the HPA's bias point, and in many cases, an RC high-pass element is used for gain shaping and stabilization. As the ISMN will be mirrored along the x-axis for the final HPA, the total width of the network (w_{tot} in Fig. 4) fixed. Also, the locations of the gate pins (shown as red dots) are fixed by the output matching network.

This basic topology can be altered by adding further components or rearranging some of them. In the authors' experience, a series line before the drain bus connections (MSL in Fig. 4) can improve matching considerably. Choosing appropriate dimensions, this topology can form a fourth-order bandpass network. However, in practical applications, a third- or second-order nonminimum network [30] is realized using the same topology since it often features lower insertion loss while still providing adequate RL performance. This is shown in Fig. 5, which shows a comparison of second- and third-order matching responses from the topology from Fig. 4.
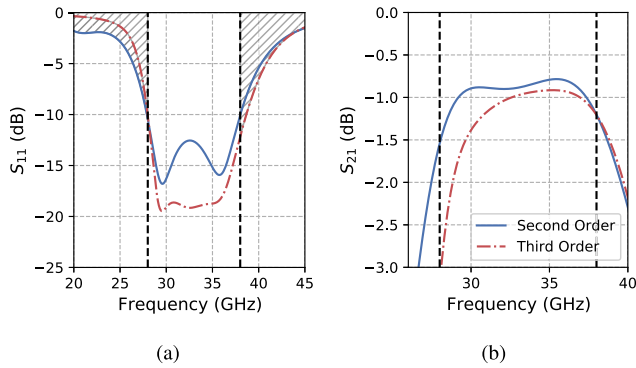
Fig. 5. Comparison of (a) S_{11} and (b) S_{21} of nonminimum second- and third-order ISMNs that are optimized to yield a minimum reflection in the band between 28 and 38 GHz (vertical lines).
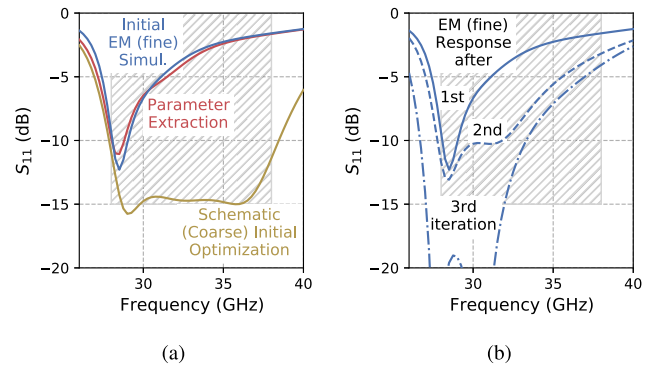

Fig. 6. Illustration of the ISMN implementation process using algorithm C. (a) Response of the schematic coarse model after initial optimization (yellow) and EM fine model response (blue) of the same structure. Note the strong disagreement between the curves—the design goal is for the coarse and fine model responses to be equal. To create the next iteration, a parameter extraction step is performed, in which the coarse model is fitted to the fine model response (red). (b) EM fine model progression for the first three iterations.

The vertical lines represent the band of interest. The entire area under the S_{11}-curve counts toward the integrals of (1) and (3). Therefore, the area outside the band of interest (shown hatched) limits the achievable in-band RL. As can be seen, the third-order network features higher insertion loss but satisfies the matching goal of $\Gamma = S_{11} < -15$ dB in the entire frequency range of interest. On the other hand, the second-order network only exhibits an $S_{11} \leq -10$ dB but features lower insertion loss. Considering the case laid out in Section III, where we assume a constant in-band reflection coefficient and total reflection out-of-band, this would only be satisfied by a network of infinite order. However, we can come reasonably close; for the third- and second-order responses, 63.3% and 51.0% of the area under the curve are in the band of interest, respectively.

For the experiment described in Section V, we parameterize the width and height of the elements highlighted in Fig. 4. The stabilization element's values are given by the active device and are thus fixed for the purpose of this experiment, as well as the RF blocking capacitors.

To describe the electrical behavior of the structure with a scalable schematic model, we use the components provided by the microstrip library in Keysight ADS wherever viable. In addition, we employ fab models where needed, for example, in case of the MIM capacitors. It is crucial to describe the network as accurately as possible. The parasitic behavior of parts, such as the T junction, the X junction, or a 45° curve, plays an important role in this and must be included from the library. Using an optimizer and the aforementioned goal of $\Gamma < -15$ dB between 28 and 38 GHz as seen by the HEMT drain, we obtain the result shown as a yellow curve in Fig. 6(a). The goal keepout area is illustrated as a hashed box in the same plot. As can be seen, the optimized schematic simulation satisfies the goal adequately in the band of interest.

V. ISMN IMPLEMENTATION STRATEGIES

This section deals with the implementation of the network developed in Section IV, i.e., finding a physical realization that matches or surpasses the predicted reflection coefficient of the schematic simulation. In the experience of the authors, this part of the design process will often be very time-consuming

and laborious, especially for layouts in which a relatively large number of transistors are to be parallelized.

To evaluate the initial guess produced by the schematic simulation, we translate the schematic to a layout and then to a 3-D model. In Fig. 6(a), the reflection coefficient Γ of the schematic simulation is plotted over frequency (yellow). The blue curve in turn shows the result of an electromagnetic (EM) simulation of the same structure. Clearly, the reflection coefficient is severely degraded and the design is in this state unacceptable for production. This is a typical effect in a complex layout like this, where extensive coupling between adjacent structures occurs. As an example, consider the parallel lines of the drain bus (*Dbus*) and the power splitter (*Sp*lit) in Fig. 4, which are only separated by the size of the MIM capacitor *Blk*.

Assuming that the EM simulation is an accurate representation of the network properties, this raises the question of how to improve the EM simulation result. An elegant way of doing this is termed *space mapping* and has been researched in the past [11]–[15], mainly for passive structures such as microstrip or waveguide filters. In [15], a summary of space-mapping approaches can be found. As the mathematical intricacies have already been derived in the literature, we do not need to repeat them here in detail. Yet, we will give a short summary to facilitate an understanding of the experiment we conducted. The basic concept of space mapping is to use a *coarse* model R_c that is computationally cheap and still implements physical knowledge of the structure in question—such as a detailed schematic. From this coarse model, a *surrogate* model R_s is deduced using a suitable transformation. In the simplest form, R_c and R_s are identical. While performing the algorithm, in each iteration, we calibrate the surrogate model from a *fine* model R_f (e.g., given by the EM simulation) to constantly decrease the error it contains. In case of convergence, the final iteration yields the design variables where the fine model represents a similar response (e.g., Γ over frequency) as the original coarse model prediction.

The advantage of this procedure is a potential decrease in the time needed to arrive at a usable layout because the fine model is only evaluated once per algorithm iteration, while the surrogate model is used for optimization. Compared to a layout optimizer, the number of EM-model evaluations is reduced by this procedure. This allows the designer to make custom changes to the layout in each algorithm iteration. For example, the procedure allows the designer to decide between passes to fold a line if it violates spatial constraints. To evaluate this approach for the design of large and broadband HPA networks with the boundary conditions mentioned in Section IV, we formulated two algorithms that are variants of basic idea of input space mapping [14]. In Sections V-A–V-G, we introduce them and propose an efficient implementation using Keysight ADS as schematic simulation tool.

A. Algorithms A and B

A flowchart of algorithms A and B is shown in Fig. 7 (top). They share the same flow but vary in the factor c. As outlined before, the algorithms require two models: a coarse and a fine model. The coarse model is a detailed schematic employing the microstrip library in ADS. It is used as the surrogate model by introducing a vector $\vec{p}_{\mathrm{cal},i}$, which contains an offset value for each design parameter. This offset can be interpreted as a calibration of the coarse model and is initially set to 0.

In each iteration, the surrogate model's offset vector is updated from the previous iteration. The newly calibrated surrogate model is then optimized toward the specified goal, in our case for the desired Γ in the band of interest. This optimization yields a set of optimized design variables $\vec{p}_{\mathrm{des},i}$ which is used to synthesize a layout. The layout, translated to an EM model, serves as the fine model. Using an EM simulator such as CST or HFSS, the fine model S-parameters are calculated and transferred back to ADS. In the next step, the surrogate model response is fitted to that of the fine model (*parameter extraction*), yielding a vector $\vec{p}_{\mathrm{extr},i}$. If the surrogate and the fine model responses line up perfectly, $\vec{p}_{\mathrm{extr},i}$ is equal to the initial surrogate optimization $\vec{p}_{\mathrm{opt},0}$. Generally,

$$\vec{p}_{\mathrm{dev},i} = \vec{p}_{\mathrm{extr},i} - \vec{p}_{\mathrm{opt},0} \tag{12}$$

represents the deviation between the initial optimized surrogate and the fine model in iteration i, expressed in the design parameter space. Using the deviation vector and the damping factor c, we can calculate the next calibration vector

$$\vec{p}_{\mathrm{cal},i} = c \cdot \sum_{n=1}^{i} \vec{p}_{\mathrm{dev},i}. \tag{13}$$

With this value, the next iteration starts—using the updated offset values from $\vec{p}_{\mathrm{cal},i}$, the surrogate model response is reoptimized, progressively improving its prediction of the fine model. The algorithm converges once the fine/surrogate model response deviation falls below a certain threshold ϵ, which can again be judged by examining the deviation value in the parameter space

$$|\vec{p}_{\mathrm{dev},i}| < \epsilon. \tag{14}$$

For algorithm A, we set the damping factor c to 1.0, whereas for algorithm B, it was set to 0.5.

B. Algorithm C

Algorithm C represents a simplified approach of Algorithms A and B and its evaluation is shown in Fig. 6. It does not make use of design parameter offsets. While in Algorithms A/B, the surrogate model is reoptimized with updated parameter offsets in each iteration, algorithm C only optimizes the surrogate model once at the start of the procedure. In the subsequent iterations, a parameter extraction [shown as a red curve in Fig. 6(a)] is performed, allowing us to calculate $\vec{p}_{\mathrm{dev},i}$ as shown in (12). The inverse of the design parameter deviation is then directly applied to the design parameter set

$$\vec{p}_{\mathrm{des},i+1} = \vec{p}_{\mathrm{des},i} - c \cdot \vec{p}_{\mathrm{dev},i}. \tag{15}$$

The updated design parameter set is then used to reevaluate the fine model. The progression of the ISMN-fine model response is shown in Fig. 6(b) for the first three iterations, showing systematic progress toward the specified goal. As in algorithms A and B, c is a damping factor reducing the impact of the changes applied to the design parameters. If its value is set too high, electrical characteristics that are unaccounted for in the surrogate model can lead to overcompensation and prevent convergence of the algorithm. For Algorithm C, c is set to 0.5. Again, convergence is judged by the similarity of extracted and original design parameters, as shown in (14).

C. Analysis

Comparing the approaches of Algorithms A/B and C, the fundamental solution found by Algorithms A/B can vary between iterations because an optimization step is performed in each iteration after parameter extraction. In contrast, Algorithm C continually tries to replicate the solution found in the initial optimization step. Therefore, Algorithms A/B may yield a better result in cases where the initial solution of the coarse model is not achievable by the fine model (e.g., due to inaccuracies of the coarse model). As an example, consider a power splitter (fork) as it occurs in the ISMN shown in Fig. 4. If the parallel microstrip lines after the fork T junction are routed very closely, their common-mode characteristic impedance Z_0 will increase considerably. In case a low Z_0 is required by the initial optimization, the solution will not be realizable in some instances. This can render algorithm C unable to solve the problem unless specific constraints are enforced in the initial optimization.

D. Practical Implementation

To be of practical use to the designer, the implementation of an MMIC design procedure is a critical factor. For the algorithms described in Section V, a reference implementation has been developed for this article. Using Keysight ADS, it has been found that Algorithm C is simpler to describe in an electrical schematic than Algorithms A/B. This is due to Algorithm C's simpler structure (see Fig. 7) in which it features an initial optimization of the coarse model, after which only a parameter extraction step is performed in each iteration.
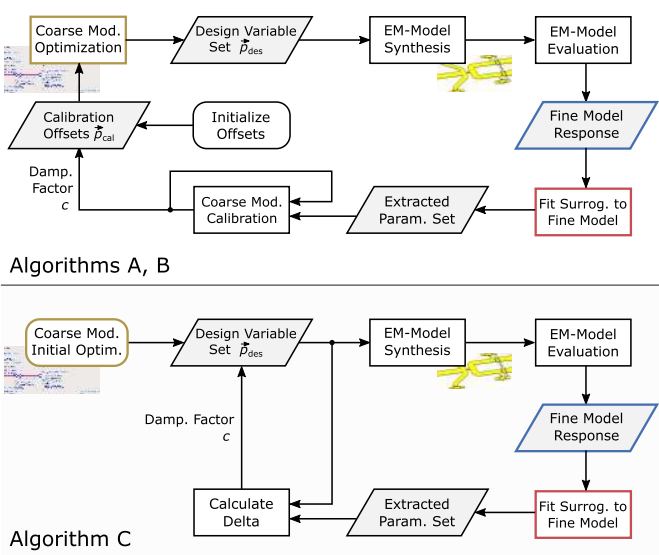
Fig. 7. Flowchart representations of the algorithms used to realize the ISMN. Algorithms A and B share the required steps but use different factors c. For A, $c = 1$, whereas for B and C, $c = 0.5$. The steps marked by colored frames correspond to the equally colored curves in Fig. 6.
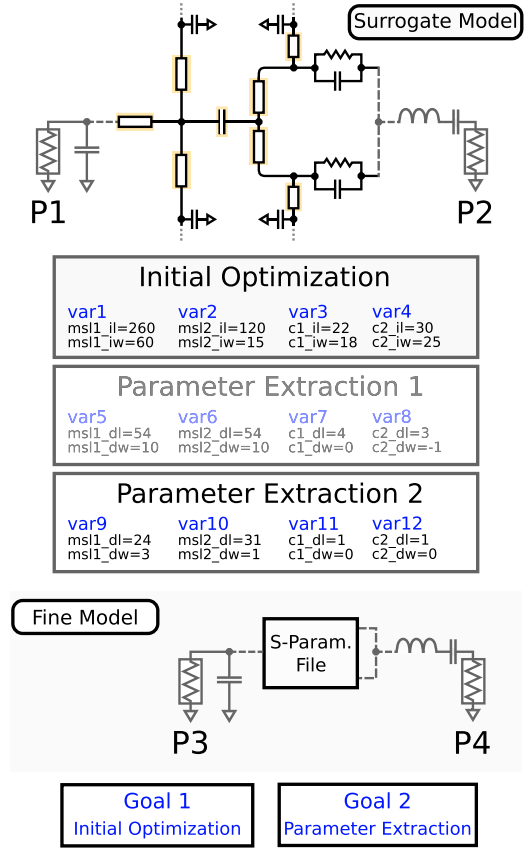

Fig. 8. Illustration of a reference implementation of algorithm C in a schematic editor such as Keysight ADS. The relevant elements of the surrogate model (top) are parameterized using the variable blocks below. Each *parameter extraction* variable block pertains to one iteration of the algorithm.

Fig. 8 shows an overview of a reference implementation of Algorithm C in Keysight ADS. It consists of three basic segments: the surrogate model, the fine model, and the simulation setup.

1) Surrogate Model: The surrogate model aims to approximate the fine model as closely as possible while remaining computationally cheap. In our case, we employ the ADS microstrip library and schematic circuits of Fraunhofer IAF's MIM capacitors. In Fig. 8, the elements highlighted in orange are changeable. In the implementation, each of these elements is parameterized using the variable blocks shown below the model. An element's value is comprised of the initial optimization and the current parameter extraction value, for example

$$\text{msl1\_l} = \text{msl1\_il} + \text{msl1\_dl}. \quad (16)$$

While the initial optimization variable block remains active throughout the execution of the algorithm, only one of the parameter extraction blocks is active at a time. For each new iteration, the previous block is copied and appended to the array.

2) Fine Model: The fine model is essentially an S-parameter block that serves to include the results obtained by the EM solver in the schematic simulation. It is connected in the same way as the surrogate model: using the FET loadline equivalent circuit on the left and the gate equivalent circuit on the right.

3) Simulation Setup: This section contains two goal setups. For the initial optimization, we use a setup to minimize the RL and/or the insertion loss of the surrogate model, i.e., the goal is set to obtain argmin abs(S_{11}) in the frequency range of interest. For the parameter extraction steps, a second setup is used where the goal is to minimize the difference between the surrogate and the fine model responses. In the schematic of Fig. 8, this means that we are looking for

$$\text{argmin abs}(S_{33} - S_{11}). \quad (17)$$

In addition, we can also include the difference in insertion loss, i.e.,

$$\text{argmin abs}(S_{43} - S_{21}). \quad (18)$$

The algorithm is then performed by first enabling the initial optimization goal and solving for the optimum parameters of the coarse model (parameter extraction block set to 0). Using an optimization controller with a random or gradient optimization in ADS, an adequate solution is usually found within 1 or 2 min. After an initial fine model evaluation, the fine model response is compared to that of the coarse model using the second goal, yielding the values of the first parameter extraction block. To prepare the fine model for its next evaluation, the parameter extraction values are multiplied with the damping factor c and subtracted from the design variable set. Next, a new fine model is generated and evaluated. This procedure is repeated until we obtain a sufficient fine model response.

E. Enforcing Physical Limitations

When realizing a closely spaced network such as the ISMN at hand, we will often find that the resulting dimensions deviate considerably from the initial schematic prediction. This can be a problem if the space available for an element is exhausted, and therefore, its size must be restricted in further algorithm

iterations. For Algorithm C, a design variable can be fixed by setting its parameter extraction variable block to 0 and excluding the value in question from further optimization runs (iterations). However, this technique is limited in some cases if the parameter extraction run cannot achieve an adequate match between the previous fine model response R_{f} and the surrogate model R_{s}. In that case, a restart of the algorithm using different constraints on the initial optimization can be in order. For Algorithms A/B, this step is simpler, as we can set boundaries as needed in the coarse model optimization step of each iteration.

F. Experimental Comparison

In order to evaluate and compare the performance of the three configurations described in the sections above, we employed them separately to realize the ISMN shown in Fig. 4. In all three cases, we used the same schematic prototype as a surrogate model after the same initial optimization as a starting point. The surrogate model initial response is used as calculated in Section IV. To judge the algorithm result progression, we utilize an alternative method to the one outlined in Section V-A. This approach is based on the keepout area shown as a hatched rectangle in Fig. 10. For each iteration, all N frequency points where the keepout area was violated were included in the following squared error sum:

$$e_i = \frac{1}{N} \sum_{n=1}^{N} (R_{\mathrm{f}}(n) - g)^2 \tag{19}$$

where g is the minimum RL specification and R_{f} is the fine model response at frequency point n.

In Fig. 9, the progression of e_i is shown for Algorithms A–C. Algorithm A exhibits a minimum error sum of $e_3 = 4$ in the third iteration, which subsequently increases again, oscillating around a value of 9 ± 5. Even with ten iterations, we could not observe convergence and thus stopped its execution.

Algorithm B, which features the same basic steps as Algorithm A but has a damping factor of $c = 0.5$, first shows a comparatively slow decay and even an increase in e_i in the first three iterations. However, in the fourth iteration, a considerable improvement can be observed, and in iteration 6, we obtain a satisfactory result of $e_6 < 0.1$.

Finally, Algorithm C (reduced form with $c = 0.5$) exhibits a relatively slow but monotonous decay of the error function up to iteration 7, where the remaining error e_7 is equal to less than 0.1.

In Fig. 10, the resulting RL curves are plotted for each of the algorithms after the final passes. Interestingly, the resulting responses R_{f} of Algorithms B and C are reasonably similar and are both an improvement on the original schematic prediction. This is somewhat unexpected at first, considering that the methodology tries to replicate the schematic response using the EM-simulated fine models. On the other hand, the microstrip library models assume the lines and junctions to be perfectly isolated from each other, which is not the case in a tight layout such as the one at hand. Therefore, some parasitic effects that
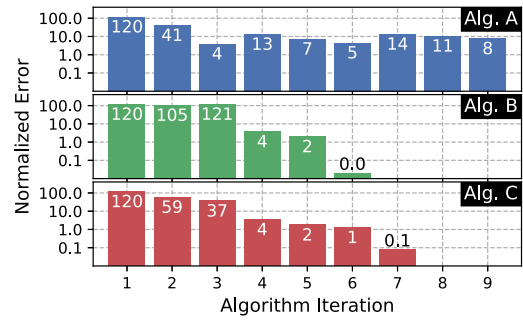

Fig. 9. Normalized square error progression of the reflection coefficient Γ for the three different algorithms while optimizing the final ISMN. The optimization goal was set to $\Gamma \leq -15\,\mathrm{dB}$ in the frequency range of 28–38 GHz. The schematic prototype of the network exhibits a squared error sum of 1.3.
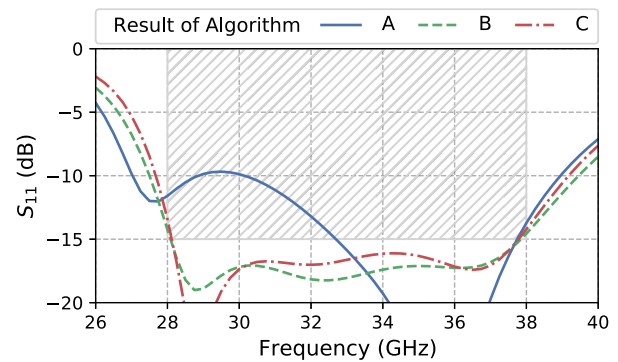

Fig. 10. ISMN input reflection coefficient Γ as seen by the drain from EM simulation after the optimization algorithms's final passes. The optimization goal of $\Gamma_{\mathrm{m}} \leq -15\,\mathrm{dB}$ is shown as a hatched rectangle.

are present in the models affect the full structure differently, permitting an improved reflection coefficient to be realized.

As mentioned before, Algorithm A did not converge and thus does not satisfy the requirements, although its result response $R_{\mathrm{f},10}$ is still an improvement on the first evaluation of the fine model $R_{\mathrm{f},1}$.

Fig. 11 shows the initial and final layouts of the ISMN elements as well as the layouts resulting from unconstrained execution of Algorithms B and C. It is noteworthy that both solutions, although similar in return and insertion loss, do not share the same dimensions. Moreover, comparing the initial layout to the final design values, we can note a substantial deviation. One example is the progression of the dimensions of series line 3, which is significantly shortened and increases in width. We can also see that the lines running in parallel (after the power splitter) are significantly reduced in length and/or increased in width. This change compensates for the coupling between them, as the impedance value of coupled lines increases for a given linewidth.

G. Result Assessment

The results indicate that the damping factor c is an important adjustment to achieve convergence. Algorithm A's fine model evaluations $R_{\mathrm{f},i}$ exhibit jumps in the matching resonances from below to above the band of interest, which is consistent
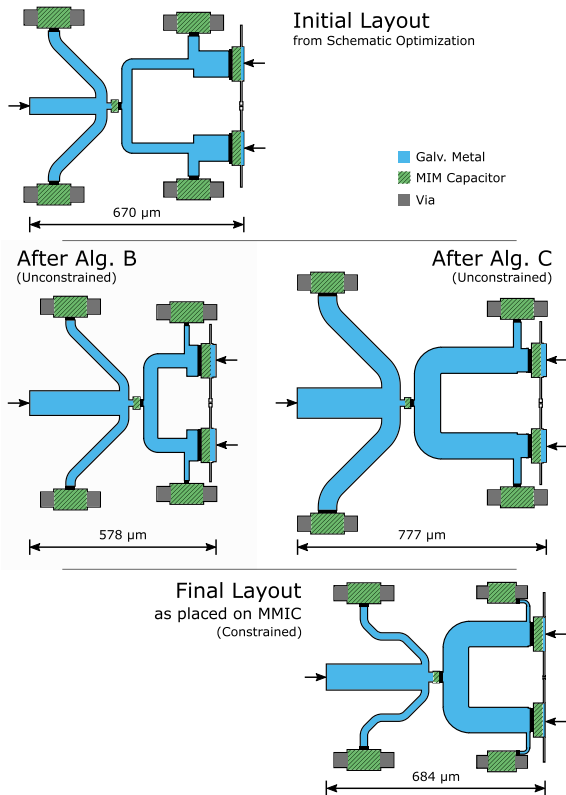
Fig. 11. Comparison of the initial layout (top) to those after performing Algorithms B and C (center) and to the final layout as placed on the PA MMIC (bottom). The final layout employed in the MMIC is subject to multiple physical and electrical constraints. Note the differences in component lengths and widths with very similar network RF-responses.

Fig. 12. Simplified schematic depiction of the developed amplifier using a mirror plane to indicate the symmetry of the structure.

with overcorrections caused by the space-mapping algorithm. A factor of $c = 0.5$ seems reasonable if the coarse model is a sufficiently accurate approximation of the structure.

In conclusion, given an adequate damping factor c, both the full (B) and the reduced algorithm C can solve the problem. However, it turns out that preconditioning of the initial guess is more important for the reduced algorithm C than for the full algorithm. More specifically, when performing the initial optimization, it is important to set boundary conditions that avoid effects not represented in the coarse model. An example for this could be thick microstrip lines that are routed closely next to each other: coupling effects between them will be nontrivial to represent in a schematic. The effect a missing representation will have can be more pronounced for the reduced algorithm as it always aims to reproduce the initial guess.

On the other hand, in the full algorithm B, the parameter extractions are used to gradually improve the calibration. In the case of an unreachable optimum, B can switch to a different local optimum. In this context, the full algorithm B is superior to the reduced one. However, the reduced algorithm is considerably easier to implement and takes less user input to perform. Compared to the simple setup as shown in Fig. 8, the full algorithm B needs an additional parameter block storing the calibration offsets for each iteration. Furthermore, the required user input is increased. However, in the experi-
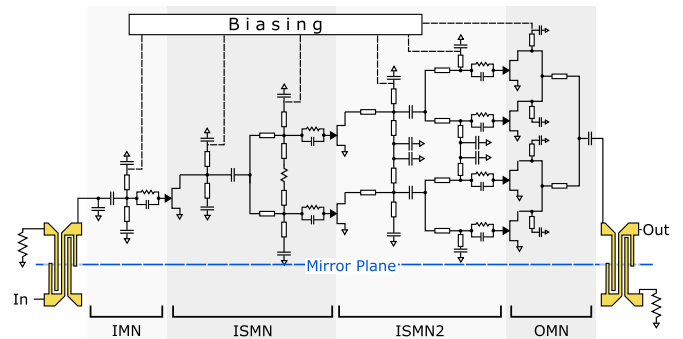
ment, it converged after six steps instead of 7 for the reduced algorithm C. Thus, depending on the complexity of the problem, the additional user input of the complex algorithm B can be warranted, for example, if the EM simulation time of additional passes is prohibitive. On the other hand, in cases with less significant EM simulation times and those where the limitations of the coarse model are well known, the reduced algorithm can be preferred.

VI. HPA Design

In order to demonstrate the usefulness of the approach outlined in the foregoing sections, we designed an HPA using IAF's 100-nm GaN-on-SiC process (see Section II for details on the technology). The MMIC is intended to be used in a large system and needs to cover most of the Ka-bands. A gain magnitude in excess of 20 dB is required. To meet these requirements, a three-stage topology was adopted featuring a staging ratio of 1:2, with eight HEMTs in the final stage. Similar to the concept in [20], eight-finger HEMTs with a unit gate width (UGW) of 60 μm each were employed, which equals a total gate width of 3.72 mm in the final stage. To ensure high RL, a balanced topology was implemented using a four-finger Lange coupler. Due to its comparatively simpler implementation, space-mapping algorithm C was used for each of the required matching networks. We found that for the ISMNs, physical limitations have to be enforced extensively (approach described in Section V-E). Some important constraints include the total network height, the HEMT port distances, and the distance between the HEMTs and the MIM capacitor vias. An overview of the ISMNs after Algorithms B/C and a comparison to the finalized (and constrained) MMIC layout is shown in Fig. 11. The ISMN MMIC layout constitutes a base cell and is mirrored along the Y-axis to create a massively parallelized IC layout. A simplified overview of the final schematic is given in Fig. 12, featuring two of the aforementioned ISMN base cells, input and output matching, and a second ISMN. It is doubled along the mirror plane (dashed line) to create the fully parallelized amplifier.

VII. Measurements and Analysis

A micrograph of the processed MMIC is shown in Fig. 13. We carried out a small-signal wafer mapping with a nominal drain voltage V_D of 15 V and a drain current density i_D
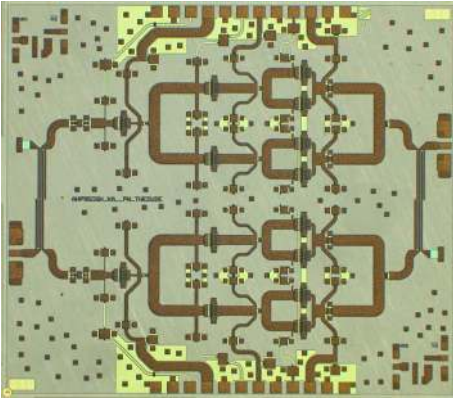
Fig. 13. Micrograph of the processed HPA. The layout's dimensions are 4 mm × 3.5 mm for a total area of 14 mm².

Fig. 15. Output power and PAE in 3-dB gain compression in the Ka-band frequency range. The MMIC was biased at a drain voltage of $V_D = 15$ V and a drain current density of $i_D = 50$ mA/mm.
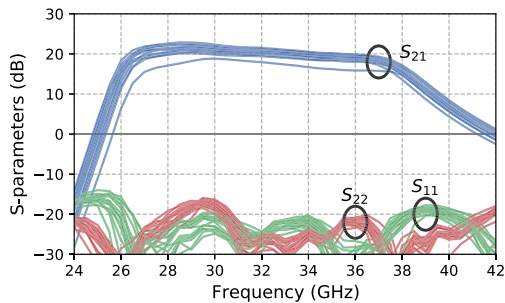

Fig. 14. Small-signal measurement data of 15 samples taken from one wafer. The amplifiers were biased at $V_D = 15$ V and $i_D = 50$ mA/mm.

Fig. 16. Relative 1-dB bandwidth over maximum output power of recent GaN HPA publications.

of 50 mA/mm. The measured S-parameters over frequency of 15 samples are shown in Fig. 14. As a result of the balanced architecture, both the input and output RL magnitudes measure below −15 dB in the entire Ka-band. This also indicates that the Lange couplers perform as designed. Furthermore, the S_{21} curves exhibit a flat characteristic, with a 3-dB small-signal band between 25.7 and 36.6 GHz (small-signal RBW of 35%). Above 37 GHz, the gain decreases at about 5 dB per GHz, which can be attributed mainly to the HEMT maximum available gain (MAG) characteristics. The large-signal frequency characterization was carried out in the same bias point.

Fig. 15 shows the output power and power-added efficiency (PAE) in 3 dB of gain compression for the entire Ka-band frequency range. The 1-dB large signal-band ranges between 28.0 and 39.0 GHz, which equates to a fractional bandwidth of RBW = 32.8%—interestingly, only slightly lower than the 3-dB small-signal band and shifted upward in frequency by about 1.5 GHz. A maximum output power of 38.2 dBm was measured at 30 GHz, the peak PAE of 26.1% at 31 GHz. Similar to the output power curve, the PAE characteristic exhibits a good flatness across the band, with a minimum of 22.4% at 39 GHz.

VIII. STATE OF THE ART

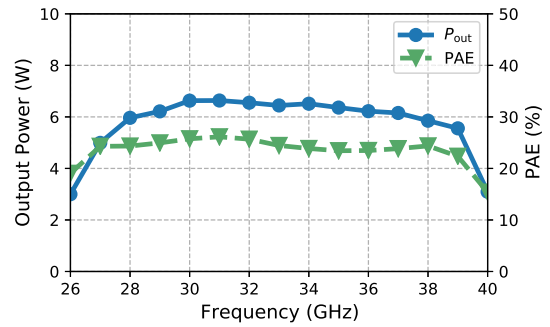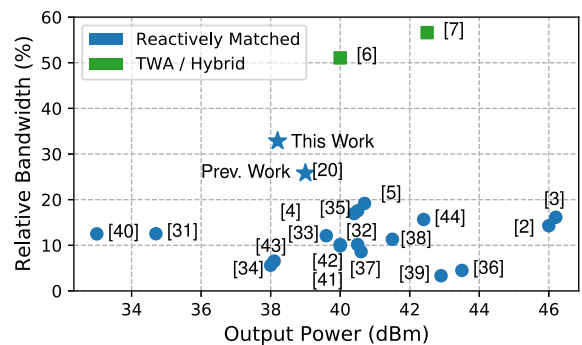A substantial amount of research has been conducted toward HPAs in the Ka-band frequency range. Fig. 16 shows an overview of recent publications in the power range above 33 dBm, with the large-signal relative 1-dB bandwidth plotted over the maximum recorded output power (see (5) for the definition of the RBW). A more detailed overview of a subset of publications is given in Table II. Note that some papers, such as [4] and [34], do not include full-band large-signal power measurements. In these cases, the available data have been used for Fig. 16. Most publications of interest describe reactively matched topologies (blue symbols), while two traveling-wave/hybrid amplifier publications were included for comparison (green symbols).

In terms of bandwidth, distributed amplifiers [6], [7] surpass reactively matched amplifiers considerably, easily exceeding 50% of RBW. On the other hand, distributed amplifiers also have disadvantages compared to reactively matched topologies, in which they require careful design of the power distribution among the circuit's transistors [6] and often reach limited efficiency [8].

With over 46 dBm, the highest output power was reported by Din et al. [2] and Roberg et al. [3]. Both of these HPAs utilize the highest drain voltage in the set ($V_D = 28$ V) and operate in the lower end of the Ka-band (below 32 GHz). Of course, a higher drain voltage often increases the loadline resistance and thus decreases the achievable bandwidth, especially for amplifiers that are limited by their drain matching bandwidth

TABLE II

SELECTION OF RECENT PUBLICATIONS OF GaN HPAs OPERATING IN Ka-BAND

$P_{\text{out,max}}$ (dBm)	Frac. BW (%)	1-dB Band (GHz)	PAE_{max} (%)	S_{21} (dB)	V_{DD} (V)	Die Area (mm^2)	Gate Length (nm)	Topology	Reference
40.0	51.1	21.0–35.4	19.0	24	20.0	15.6	150	TWA	[6]
42.5	56.6	19.0–34.0	27.0	23	20.0	9.0	150	Hybrid	[7]
40.4	16.9	27.0–32.0	27.5	32	20.0	11.7	150	Reactive	[4]
46.0	14.3	26.0–30.0	36.0	22	28.0	13.5	200	Reactive	[2]
43.4	4.5	26.1–27.3	21.0	12	24.0	23.6	150	Reactive	[36]
40.6	8.6	30.0–32.7	35.0	28	13.0	15.8	100	Reactive	[37]
40.7	19.2	29.7–36.0	34.0	25	22.5	13.4	150	Reactive	[5]
46.2	16.1	26.2–30.8	25.0	25	28.0	17.6	150	Reactive	[3]
33.0	12.5	37.5–42.5	34.0	22	8.0	11.0	100	Reactive	[40]
39.0	25.8	27.0–35.0	28.0	21	15.0	15.8	100	Reactive	[20]
38.2	32.8	28.0–39.0	26.1	20	15.0	14.0	100	Reactive	This work

Note: In column "Frac. BW", the fractional large-signal 1-dB bandwidth is shown.

(see Section III). Typically, the considered publications report an RBW of between 8% and 20%, with some exceptions with amplifiers optimized toward single-frequency operation [34], [39]. Compared to this trend, with an RBW of 32.8%, the HPA introduced in this work exhibits the highest relative large-signal bandwidth of all reactively matched amplifiers.

IX. CONCLUSION

In this article, we summarize the theoretical matching boundaries and show the limitations they impose on real-world amplifier design. Starting with a common schematic prototype, we investigate the question of how to realize its electrical response in a densely routed, massively parallelized layout. To that end, we develop a comprehensive study on the application of space-mapping techniques toward the design of HPAs. We derive three reference design procedures and compare their performance in terms of convergence, speed, and practicality when laying out a densely routed HPA ISMN. Subsequently, we demonstrate the usefulness of the study by designing the networks of a compact three-stage eight-way wideband HPA in the Ka-band. The processed MMIC features a 1-dB large-signal bandwidth of more than 11 GHz and thus covers most of the Ka-band with an output power exceeding 6 W in 3 dB of gain compression. This demonstrates the highest combination of power and bandwidth achieved to date using a reactively matched topology in the Ka-band, suggesting that the method developed in Section V is a useful tool to increase the bandwidth of a circuit or, more generally, to reproduce the electrical characteristic of a prototype network with high accuracy.

ACKNOWLEDGMENT

REFERENCES

[1] W. Q. Lohmeyer, R. J. Aniceto, and K. L. Cahoy, "Communication satellite power amplifiers: Current and future SSPA and TWTA technologies," *Int. J. Satell. Commun. Netw.*, vol. 34, no. 2, pp. 95–113, Mar. 2016.

[2] S. Din, M. Wojtowicz, and M. Siddiqui, "High power and high efficiency Ka band power amplifier," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2015, pp. 1–4.

[3] M. Roberg, T. R. Mya Kywe, M. Irvine, O. Marrufo, and S. Nayak, "40 W Ka-band single and dual output GaN MMIC power amplifiers on SiC," in *Proc. IEEE BiCMOS Compound Semiconductor Integr. Circuits Technol. Symp. (BCICTS)*, Oct. 2018, pp. 140–143.

[4] C. F. Campbell, Y. Liu, M.-Y. Kao, and S. Nayak, "High efficiency Ka-band gallium nitride power amplifier MMICs," in *Proc. IEEE Int. Conf. Microw., Commun., Antennas Electron. Syst. (COMCAS)*, Oct. 2013, pp. 1–5.

[5] V. D. Giacomo-Brunel *et al.*, "Industrial 0.15-μm AlGaN/GaN on SiC technology for applications up to Ka band," in *Proc. Eur. Microw. Integr. Circuits Conf.*, Sep. 2018, pp. 13–16.

[6] C. F. Campbell, S. Nayak, M.-Y. Kao, and S. Chen, "Design and performance of 16–40 GHz GaN distributed power amplifier MMICs utilizing an advanced 0.15μm GaN process," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2016, pp. 10–13.

[7] C.-H. Han and H.-Q. Tao, "A 18-40 GHz 10 W GaN power amplifier MMIC utilizing combination of the distributed and reactive matching topology," in *Proc. 14th Eur. Microw. Integr. Circuits Conf. (EuMIC)*, Sep. 2019, pp. 228–231.

[8] J. Kim and Y. Kwon, "A high-performance GaN-modified nonuniform distributed power amplifier," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 5, pp. 1729–1740, May 2020.

[9] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. New York, NY, USA: D. Van Nostrand Co., 1945.

[10] R. M. Fano, "Theoretical limitations on the broad-band matching of arbitrary impedances 1," *J. Franklin Inst.*, vol. 249, no. 1, pp. 57–83, 1950.

[11] J. W. Bandler, R. M. Biernacki, S. H. Chen, and D. Omeragić, "Space mapping optimization of waveguide filters using finite element and mode-matching electromagnetic simulators," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 1, pp. 54–70, Jan. 1999.

[12] J. W. Bandler *et al.*, "Space mapping: The state of the art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 337–361, Jan. 2004.

[13] J. W. Bandler, Q. S. Cheng, D. M. Hailu, and N. K. Nikolova, "A space-mapping design framework," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 11, pp. 2601–2610, Nov. 2004.

[14] S. Koziel, J. W. Bandler, and K. Madsen, "A space-mapping framework for engineering optimization–theory and implementation," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 10, pp. 3721–3730, Oct. 2006.

[15] S. Koziel, Q. S. Cheng, and J. W. Bandler, "Space mapping," *IEEE Microw. Mag.*, vol. 9, no. 6, pp. 105–122, Dec. 2008.

[16] S. Amari, C. LeDrew, and W. Menzel, "Space-mapping optimization of planar coupled-resonator microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 5, pp. 2153–2159, May 2006.

[17] F. Colomb and A. Platzker, "A 3-watt Q-band GaAs pHEMT power amplifier MMIC for high temperature operation," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2006, pp. 897–900.

[18] D. Schwantuschke, P. Bruckner, S. Wagner, M. Dammann, M. Mikulla, and R. Quay, "Enhanced GaN HEMT technology for E-band power amplifier MMICs with 1W output power," in *Proc. IEEE Asia Pacific Microw. Conf. (APMC)*, Nov. 2017, pp. 395–398.

[19] M. Cwiklinski et al., "D-band and G-Band high-performance GaN power amplifier MMICs," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 12, pp. 5080–5089, Dec. 2019.

[20] P. Neininger, L. John, P. Bruckner, C. Friesicke, R. Quay, and T. Zwick, "Design, analysis and evaluation of a broadband high-power amplifier for Ka-band frequencies," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2019, pp. 564–567.

[21] M. R. Duffy, G. Lasser, G. Nevett, M. Roberg, and Z. Popovic, "A three-stage 18.5–24-GHz GaN-on-SiC 4 W 40% efficient MMIC PA," *IEEE J. Solid-State Circuits*, vol. 54, no. 9, pp. 2402–2410, Sep. 2019.

[22] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA, USA: Artech House, 2006.

[23] D. M. Pozar, *Microwave Engineering*. Hoboken, NJ, USA: Wiley, 2009.

[24] A. Grebennikov, N. Kumar, and B. S. Yarman, *Broadband RF and Microwave Amplifiers*. Boca Raton, FL, USA: CRC Press, 2015.

[25] C. F. Campbell, "Gallium nitride power MMICs–promise and problems," in *Proc. Integr. Nonlinear Microw. Millimetre-Wave Circuits Workshop (INMMiC)*, no. 1, Oct. 2015, pp. 12–17.

[26] F. Thome, S. Maroldt, and O. Ambacher, "Prospects and limitations of stacked-FET approaches for enhanced output power in voltage-controlled oscillators," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 3, pp. 836–846, Mar. 2016.

[27] H. Carlin, "Synthesis techniques for gain-bandwidth optimization in passive transducers," *Proc. IRE*, vol. 48, no. 10, pp. 1705–1714, Oct. 1960.

[28] H. Carlin and B. Yarman, "The double matching problem: Analytic and real frequency solutions," *IEEE Trans. Circuits Syst.*, vol. CS-30, no. 1, pp. 15–28, Jan. 1983.

[29] B. S. Yarman, "Modern approaches to broadband matching problems," *IEE Proc. H-Microw., Antennas Propag.*, vol. 132, no. 2, pp. 87–92, Apr. 1985.

[30] M. W. Medley, *Microwave and RF Circuits: Analysis, Synthesis and Design*. Norwood, MA, USA: Artech House, 1992.

[31] K. Kong, M.-Y. Kao, and S. Nayak, "Miniaturization of Ka-band high power amplifier by 0.15 μm GaN MMIC technology," in *Proc. IEEE Compound Semiconductor Integr. Circuit Symp. (CSICS)*, Oct. 2014, pp. 1–4.

[32] X. Yu, W. Hong, W. Wang, H. Tao, and C. Ren, "A millimeter wave 11W GaN MMIC power amplifier," in *Proc. 3rd Asia–Pacific Conf. Antennas Propag.*, Jul. 2014, pp. 1342–1344.

[33] Y. Noh, Y.-H. Choi, and I. Yom, "Ka-band GaN power amplifier MMIC chipset for satellite and 5G cellular communications," in *Proc. IEEE 4th Asia–Pacific Conf. Antennas Propag. (APCAP)*, Jun. 2015, pp. 453–456.

[34] P. Blount, S. Huettner, and B. Cannon, "A high efficiency, Ka-band pulsed gallium nitride power amplifier for radar applications," in *Proc. IEEE Compound Semiconductor Integr. Circuit Symp. (CSICS)*, Oct. 2016, pp. 1–4.

[35] S. Chen, S. Nayak, C. Campbell, and E. Reese, "High efficiency 5W/10W 32–38 GHz power amplifier MMICs utilizing advanced 0.15μm GaN HEMT technology," in *Proc. IEEE Compound Semiconductor Integr. Circuit Symp. (CSICS)*, Oct. 2016, pp. 32–35.

[36] Y. Yamaguchi, J. Kamioka, M. Hangai, S. Shinjo, and K. Yamanaka, "A CW 20 W Ka-band GaN high power MMIC amplifier with a gate pitch designed by using one-finger large signal models," in *Proc. IEEE Compound Semiconductor Integr. Circuit Symp. (CSICS)*, Oct. 2017, pp. 1–4.

[37] A. Gasmi et al., "10 W power amplifier and 3 W transmit/receive module with 3 dB NF in Ka band using a 100 nm GaN/Si process," in *Proc. IEEE Compound Semiconductor Integr. Circuit Symp. (CSICS)*, Oct. 2017, pp. 1–4.

[38] J. Moron, R. Leblanc, F. Lecourt, and P. Frijlink, "12 W, 30% PAE, 40 GHz power amplifier MMIC using a commercially available GaN/Si process," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2018, pp. 1457–1460.

[39] R. Welker, S. Ozev, and J. Kitchen, "Incorporating RF test measurements for efficient design flow of GaN-based power amplifiers," in *Proc. IEEE Top. Conf. RF/Microw. Power Modeling Radio Wireless Appl. (PAWR)*, Jan. 2018, pp. 69–71.

[40] F. Costanzo, R. Giofre, G. Polli, A. Salvucci, and E. Limiti, "A Q-band MMIC power amplifier in GaN on Si technology for space applications," in *Proc. Int. Workshop Integr. Nonlinear Microw. Millimetre-Wave Circuits (INMMIC)*, Jul. 2018, pp. 1–3.

[41] N. Estella, E. Camargo, J. Schellenberg, and L. Bui, "High-efficiency, Ka-band GaN power amplifiers," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2019, pp. 568–571.

[42] C. Potier et al., "10W Ka band MMIC power amplifiers based on InAlGaN/GaN HEMT technology," in *Proc. 49th Eur. Microw. Conf. (EuMC)*, Oct. 2019, pp. 270–273.

[43] B. Schmukler et al., "A high efficiency, Ka-band, GaN-on-SiC MMIC with low compression," in *Proc. IEEE BiCMOS Compound Semiconductor Integr. Circuits Technol. Symp. (BCICTS)*, Nov. 2019, pp. 38–41.

[44] K. Nakatani, Y. Yamaguchi, M. Hangai, and S. Shinjo, "A Ka-band CW 15.5 W 15.6% fractional bandwidth GaN power amplifier MMIC using wideband BPF inter-stage matching network," in *Proc. IEEE BiCMOS Compound Semiconductor Integr. Circuits Technol. Symp. (BCICTS)*, Nov. 2019, pp. 1–4.

Philipp Neininger (Student Member, IEEE) received the B.S. and M.S. degrees in electrical engineering and information technologies from the Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree at the Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), Freiburg im Breisgau, Germany.

His research interests include millimeter-wave (mm-wave) power combining and broadband designs in high-power-density composite semiconductor processes for Ka-band and above.

Laurenz John received the Master of Science (M.S.E.E.) degree in electrical engineering and information technologies from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2016.

Since 2016, he has been involved in the design and characterization of InGaAs-channel high-electron-mobility transistor (HEMT) devices and integrated circuits on GaAs and Si substrates for millimeter-wave (mm-wave) and sub-mm-wave frequency applications. He is currently a Research Engineer and the Project Leader with the Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), Freiburg im Breisgau, Germany. His current research interests include integrated circuit (IC) and power amplifier (PA) design for radar, and communication applications above 100 GHz.

Fabian Thome received the Dipl.Ing. degree in electrical engineering and information technologies from the Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2011, and the Ph.D. degree in semiconductor devices and wireless communication from the University of Freiburg, Freiburg im Breisgau, Germany, in 2020.

In 2010, he joined the Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), Freiburg im Breisgau, where he has been a Researcher and the Project Manager since 2012. His current research interests include the development of InGaAs and GaN high-electron-mobility transistor (HEMT) technologies as well as the design and characterization of linear and nonlinear semiconductor devices and monolithic microwave integrated circuits from the microwave to submillimeter-wave frequency range for applications in radio astronomy, wireless communications, and quantum computing.

Christian Friesicke (Member, IEEE) was born in Berlin, Germany, in 1981. He received the Dipl.Ing. degree in electrical engineering from the Technische Universität Hamburg–Harburg, Hamburg, Germany, in 2008.

From 2005 to 2006, he was a Visiting Scholar with the University of California at Berkeley, Berkeley, CA, USA. From 2008 to 2015, he was with the Institut für Hochfrequenztechnik, Technische Universität Hamburg–Harburg. In 2015, he joined the Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), Freiburg im Breisgau, Germany, where he is involved in monolithic microwave integrated circuit (MMIC) design activities. His current research interests are the theory and design of power amplifiers in III–V and GaN technologies.

Peter Brückner received the Diploma degree in electrical engineering and the Ph.D. degree from the University of Ulm, Ulm, Germany, in 2004 and 2008, respectively.

He is currently with the Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), Freiburg im Breisgau, Germany, where he is involved in the development of GaN HEMT technologies.

Rüdiger Quay (Senior Member, IEEE) received the Diploma degree in physics from Rheinisch-Westfälische Technische Hochschule (RWTH), Aachen, Germany, in 1997, the Ph.D. degree (Hons.) in technical sciences from the Technische Universität Wien, Vienna, Austria, in 2001, the second Diploma degree in economics in 2003, and the Venia Legendi (Habilitation) in microelectronics from the Technische Universität Wien in 2009.

In 2001, he joined the Fraunhofer Institute of Applied Solid-State Physics (Fraunhofer IAF), Freiburg im Breisgau, Germany, in various positions. He is currently the Deputy Director of Fraunhofer IAF, where he is responsible for the business fields. Since 2020, he has been a Fritz-Hüttinger Professor with the Institute for sustainable systems, Albert-Ludwig University, Freiburg im Breisgau. He has authored or coauthored over 300 refereed publications, three monographs, and contributions to two further.

Thomas Zwick (Fellow, IEEE) received the Dipl.Ing. (M.S.E.E.) and the Dr.Ing. (Ph.D.E.E.) degrees from the Universität Karlsruhe (TH), Karlsruhe, Germany, in 1994 and 1999, respectively.

From 1994 to 2001, he was a Research Assistant with the Institut für Höchstfrequenztechnik und Elektronik (IHE), TH. In February 2001, he joined the IBM T. J. Watson Research Center, IBM, Yorktown Heights, NY, USA, as a Research Staff Member. From October 2004 to September 2007, he was with Siemens AG, Lindau, Germany. During this period, he managed the RF development team for automotive radars. In October 2007, he became a Full Professor at the Karlsruhe Institute of Technology (KIT), Karlsruhe. He is currently the Director of the Institute of Radio Frequency Engineering and Electronics (IHE), KIT. He is a coeditor of three books, an author or a coauthor of 120 journal articles, and over 400 contributions at international conferences. He holds 15 granted patents. His research topics include wave propagation, stochastic channel modeling, channel measurement techniques, material measurements, microwave techniques, millimeter-wave antenna design, wireless communication, and radar system design.

Dr. Zwick has been a member of acatech (German National Academy of Science and Engineering) since 2019. His research team received over ten best paper awards on international conferences. He served on the technical program committee (TPC) of several scientific conferences. He was the General Chair of the International Workshop on Antenna Technology (iWAT 2013), Karlsruhe, in 2013, and the IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM), Heidelberg, in 2015. He was also the TPC Chair of the European Microwave Conference (EuMC) in 2013 and the General TPC Chair of the European Microwave Week (EuMW) in 2017. In 2023, he will be the General Chair of EuMW, Berlin. From 2008 until 2015, he was the President of the Institute for Microwaves and Antennas (IMA). He selected as a Distinguished IEEE Microwave Lecturer for the term 2013–2015 with his lecture on "QFN Based Packaging Concepts for Millimeter-Wave Transceivers." Since 2017, he has been a member of the Heidelberg Academy of Sciences and Humanities. In 2019, he became the Editor in Chief of the IEEE MICROWAVE AND WIRELESS COMPONENTS LETTERS.