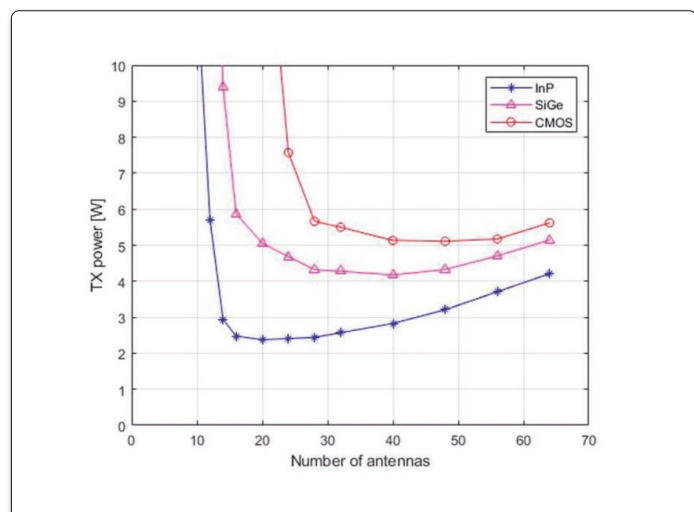


Für die drahtlose 5G- und 6G-Kommunikation

Hochskalierung von GaN- und InP-basierten Technologien

Vergleich des Stromverbrauchs von CMOS-, SiGe- und InP-Bauelementen in Senderarchitekturen in Abhängigkeit von der Anzahl der Antennen (wie auf der IEDM 2022 vorgestellt)



Es wird über Fortschritte bei der Hochskalierung von GaN-HEMTs und InP-HBTs auf eine Siliziumplattform und bei der Ko-Integration mit CMOS-Komponenten berichtet und damit über Voraussetzungen für die drahtlose Kommunikation der nächsten Generation mit hoher Kapazität.

Jahr für Jahr werden immer mehr Daten von einer immer größeren Gruppe von Nutzern drahtlos übertragen. Um mit diesem Trend Schritt zu halten und die Datenübertragung schneller und effizienter zu gestalten, wird derzeit die fünfte Generation (5G) der mobilen Kommunikation eingeführt, und die Branche blickt bereits auf die Zukunft. Während 5G Spitzen-Datenraten von 10 Gbit/s ermöglicht, wird 6G voraussichtlich ab 2030 mit 100 Gbit/s arbeiten. Neben der Bewältigung von mehr Daten und Verbindungen untersuchen die Forscher auch, wie die nächste Generation der drahtlosen Kommunikation neue Anwendungsfälle wie autonomes Fahren und holografische Präsenz unterstützen kann.

Autorin:
Nadine Collaert
Programmlinieiterin für
fortgeschrittene RF bei
Imec
www.imec-int.com

Um diese extrem hohen Datenübertragungsraten zu ermöglichen, hat die Telekommunikationsbranche die Frequenzen der Funksignale erhöht. Während 5G zunächst Frequenzbänder unter 6 GHz nutzt, werden bereits Produkte vorgestellt, die auf 28/39 GHz ausgerichtet sind. Darüber hinaus besteht ein wachsendes Interesse an der Nutzung von FR3-Frequenzbändern (6...20 GHz) für 5G-Netze, da diese ein Gleichgewicht zwischen Abdeckung und Kapazität bieten. Für 6G werden Frequenzen über 100 GHz diskutiert.

Der Wechsel zu höheren Frequenzen hat mehrere Vorteile

Es können neue Frequenzbänder genutzt werden, wodurch das Problem der Frequenzknappheit in den bestehenden Bändern gelöst wird. Je höher die Betriebsfrequenz, desto einfacher ist es, größere Bandbreiten zu erzielen. Frequenzen über 100 GHz und Bandbreiten bis zu 30 GHz ermöglichen es den Telekommunikationsbetreibern im Prinzip, in den drahtlosen Datenverbindungen Modulationsschemata niedrigerer Ordnung zu verwenden, was den Stromverbrauch senkt. Höhere Frequenzen sind auch mit kleineren Wellenlängen (λ) verbunden. Da die Größe der Antennengruppen mit λ^2 skaliert, können die Antennengruppen dichter gepackt werden. Dies trägt zu einem besseren Beamforming bei, einer Technik, die sicherstellt, dass ein größerer Teil der übertragenen Energie den vorgesehenen Empfänger erreicht.

Die Einführung höherer Frequenzen hat ihren Preis

Heute ist CMOS die bevorzugte Technologie für den Bau der kritischen Komponenten von Sendern und Empfängern. Dazu gehören die Leistungsverstärker in den Front-End-Modulen, die die Hochfrequenzsignale zu und von den Antennen senden. Je höher die Betriebsfrequenz ist, desto mehr haben die CMOS-basierten Leistungsverstärker Schwierigkeiten, die erforderliche Ausgangsleistung mit ausreichend hohem Wirkungsgrad zu liefern.

Und hier kommen Technologien wie GaN und InP ins Spiel. Aufgrund ihrer hervorragenden Materialeigenschaften sind diese III/V-Halbleiter eher in der Lage, die erforderliche Ausgangsleistung und Effizienz bei hohen Betriebsfrequenzen zu liefern. GaN hat zum Beispiel eine hohe Stromdichte, eine hohe Elektronenbeweglichkeit und eine große Durchbruchspannung. Die hohe Leistungsdichte ermöglicht auch einen kleinen Formfaktor und damit eine Verringerung der Gesamtgröße des Systems bei gleicher Leistung.

GaN und InP übertreffen CMOS bei höheren Frequenzen

In einer Modellierungsübung verglichen die Forscher am imec die Leistung von drei verschiedenen Leistungsverstärker-Implementierungen bei 140 GHz Betriebsfrequenz: eine vollständige CMOS-Implementierung, ein CMOS-Beamformer mit SiGe-Heterojunction-Bipolartransistor (HBT) und ein InP-HBT. InP war der klare Sieger in Bezug auf die Ausgangsleistung

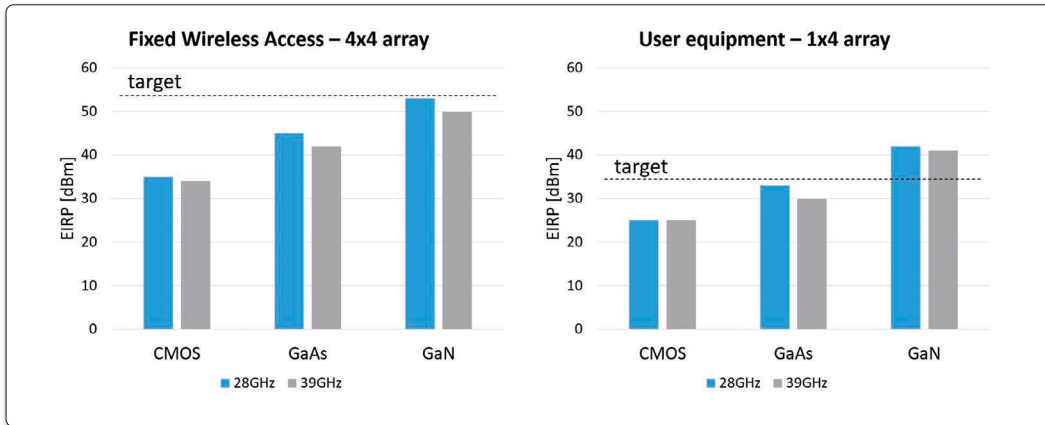


Bild 2: Ausgangsleistung für 28 GHz und 39 GHz Betriebsfrequenzen in (links) FWA- und (rechts) Nutzergeräten: ein Vergleich von drei verschiedenen Technologien (wie auf der IEDM 2022 vorgestellt)

(über 20 dBm) und die Energieeffizienz (20 bis 30%). Die Modellierungsergebnisse zeigen auch, dass bei InP der optimale Punkt für die Energieeffizienz mit einer relativ geringen Anzahl von Antennen erreicht wird. Dies ist besonders interessant für Anwendungsfälle mit geringem Platzbedarf wie Benutzergeräte (z.B. mobile Geräte).

Bei niedrigeren mm-Wellenfrequenzen zeigt GaN jedoch eine hervorragende Leistung. Sowohl bei 28 als auch bei 39 GHz wird beobachtet, dass Transistoren mit hoher Elektronenbeweglichkeit (HEMTs) aus GaN auf Siliziumkarbid (GaN-on-SiC) CMOS-basierte Bauelemente und GaAs-HEMTs in Bezug auf Ausgangsleistung und Energieeffizienz übertreffen. Es wurden zwei verschiedene Anwendungsfälle betrachtet, d. h. Fixed Wireless Access (FWA, mit 16 Antennen) und Benutzergeräte (mit vier Antennen).

Chancen und Herausforderungen für die Ausweitung des Angebots

Betrachtet man jedoch die Kosten und die einfache Integration, so können GaN- und InP-Bauelementetechnologien noch nicht vollständig mit CMOS-basierten Technologien konkurrieren. Die III/V-Bauelemente werden in der Regel auf kleinen und kostspieligen Nicht-Si-Substraten hergestellt und sind für die Herstellung großer Stückzahlen weniger geeignet. Die Integration dieser Bauelemente auf 200- oder 300-mm-Si-Wafern ist ein interessanter Ansatz, um eine

Gesamtoptimierung zu erreichen und gleichzeitig eine hervorragende HF-Leistung beizubehalten. Si-Substrate sind nicht nur billiger, sondern der CMOS-kompatible Prozess ermöglicht auch die Herstellung in großem Maßstab.

Die Integration von GaN und InP auf einer Si-Plattform erfordert eine Kombination aus neuen Transistor- und Schaltungsdesignansätzen, Materialien und Fertigungstechniken. Eine der größten Herausforderungen ist die große Gitterfehlpassung: 8% bei InP und 17% bei GaN. Dies führt bekanntermaßen zu zahlreichen Defekten in den Schichten, die letztlich die Leistung der Bauelemente beeinträchtigen.

Darüber hinaus müssen wir die GaN-on-Si- und InP-on-

Si-basierten Komponenten mit CMOS-basierten Komponenten zu einem Gesamtsystem integrieren. GaN- und InP-Technologien werden zunächst für die Realisierung der Leistungsverstärker in den Front-End-Modulen eingesetzt. Auch rauscharme Verstärker und Schalter könnten möglicherweise von den einzigartigen Eigenschaften dieser Verbindungshalbleiter profitieren. Für die Kalibrierung, die Steuerung und die Strahlformung wird jedoch weiterhin CMOS benötigt.

Im Rahmen seines Advanced RF-Programms untersucht das imec zusammen mit seinen Industriepartnern verschiedene Ansätze zur Integration von GaN- und InP-Bauelementen auf großformatigen Si-Wafern und zur Ermöglichung ihrer

heterogenen Integration mit CMOS-Komponenten. Die Vor- und Nachteile werden für verschiedene Anwendungsfälle bewertet – sowohl für die Infrastruktur (z.B. FWA) als auch für Nutzergeräte.

Verbesserung der GaN-on-Si-Technologie für RF-Leistung

Je nach Ausgangssubstrat gibt es verschiedene Varianten der GaN-Technologie: GaN-Bulk-Substrate, GaN-on-SiC und GaN-on-Si. Heute wird GaN-on-SiC weithin erforscht und bereits für Infrastrukturanwendungen, einschließlich 5G-Basisstationen, verwendet. GaN-on-SiC ist kosteneffizienter als die GaN-Basistechnologie, und SiC ist ein hervorragender Wärmeleiter, der dazu beiträgt, die bei Hochleistungs-Infrastrukturanwendungen

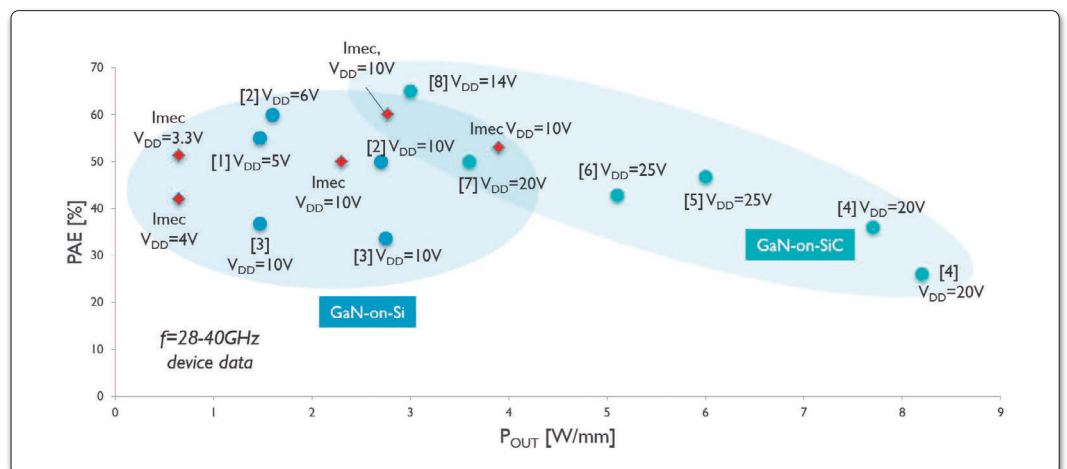


Bild 3: GaN-on-Si-Benchmarking-Daten. Die roten Daten von imec gehören zu den besten, die für GaN-on-Si-Bauelemente berichtet wurden, und sind vergleichbar mit GaN-on-SiC-Substraten (wie auf der IEDM 2022 vorgestellt). Referenzen: [1] H.W. Then et al., IEDM 2020; [2] H.W. Then et al., IEDM 2021; [3] W. Wang et al., J-EDS 2018; [4] Y.C. Lin et al., Micromachines 2020; [5] M. Mi et al., TED 2017; [6] Y. Zhang et al., EDL 2018; [7] K. Harrouche et al., HAL open science, 2020; [8] J.-S. Moon et al., MTTs 2019.

entstehende Wärme abzuleiten. Die Kosten und die begrenzte Größe des Substrats machen es jedoch weniger geeignet für die Massenproduktion.

GaN-on-Si hingegen hat das Potenzial, auf 200- und sogar 300-mm-Wafers hochskaliert zu werden. Dank jahrelanger Innovationen für Anwendungen in der Leistungselektronik hat die Integration von GaN auf großformatigen Si-Substraten enorme Fortschritte gemacht. Es sind jedoch noch weitere Verbesserungen erforderlich, um die GaN-on-Si-Technologie für eine optimale HF-Leistung vorzubereiten. Die größten Herausforderungen bestehen darin, eine vergleichbare Leistung bei großen Signalen und Zuverlässigkeit wie bei GaN-on-SiC zu erreichen und die Betriebsfrequenz zu erhöhen. Dies erfordert kontinuierliche Innovationen beim Materialstapel-Design und bei der Materialauswahl, die Reduzierung der Gate-Länge der HEMTs, die Unterdrückung von

Parasiten und eine möglichst geringe HF-Dispersion.

Imec's GaN-on-Si Prozessablauf für RF beginnt mit dem Wachstum (durch metallorganische chemische Gasphasenabscheidung (MOCVD)) einer epitaktischen Struktur auf 200-mm-Si-Wafers. Diese Struktur besteht aus einer proprietären GaN/AlGaN-Pufferstruktur, einem GaN-Kanal, einem AlN-Abstandshalter und einer AlGaN-Barriere. GaN-HEMT-Bauelemente mit TiN-Schottky-Metall-Gate werden anschließend mit einem (Niedrigtemperatur)-3-Stufen-Cu-Backend-of-Line-Prozess integriert.

Kürzlich wurden auf der GaN-on-Si-Plattform von imec wettbewerbsfähige Ergebnisse erzielt, die die Ausgangsleistung und die Leistungszusatzeffizienz (PAE) zum ersten Mal näher an die der GaN-on-SiC-Technologie heranbringen. Die PAE ist eine gängige Kennzahl zur Bewertung der Effizienz eines Leistungsverstärkers, die die

Auswirkung der Verstärkung des Verstärkers auf seine Gesamteffizienz berücksichtigt.

Die Ergänzung der Technologieentwicklung durch Modellierungsaktivitäten wird letztendlich zu einer noch besseren Leistung und Zuverlässigkeit beitragen. Auf der IEDM 2022 stellte imec beispielsweise ein Simulations-Framework vor, mit dem der Wärmetransport in HF-Bauteilen besser vorhergesagt werden kann. In einer Fallstudie mit GaN-on-Si-HEMTs ergaben die Simulationen Temperaturspitzen, die bis zu dreimal höher waren als zuvor vorhergesagt. Modellierungsarbeiten wie diese bieten weitere Anhaltspunkte für die Optimierung von HF-Bauelementen und deren Layouts in einer frühen Entwicklungsphase.

InP-on-Si für 6G-Sub-THz-Frequenzen: drei Fertigungsansätze

Wie bereits gezeigt, bieten InP-HBTs bei einer Betriebsfrequenz von 140 GHz das beste Verhältnis zwischen Ausgangsleistung

und Wirkungsgrad aller Technologien. Die Forscher wissen auch, wie man InP-HBTs für eine optimale HF-Leistung entwickelt. Die Herstellung erfolgt jedoch in der Regel auf kleinen (InP-)Substratwafers (<150 mm) und mit laborähnlichen Prozessen, die nicht CMOS-kompatibel sind.

Doch was geschieht mit der Leistung, wenn wir InP auf Si integrieren?

Die Abscheidung von InP auf Si führt bekanntermaßen zu zahlreichen Defekten, vor allem zu Fadenversetzungen und flächigen Defekten. Diese Defekte führen zu Leckströmen, die die Leistung der Bauelemente drastisch verschlechtern oder Probleme mit der Zuverlässigkeit verursachen können.

Für die Hochskalierung werden drei Ansätze in Betracht gezogen. Zwei davon beruhen auf dem direkten Wachstum von InP auf Si, ein weiterer auf der Rekonstruktion von Wafers. Alle drei Ansätze sollen kostengünstigere Lösungen bieten als die derzeitigen Technologien, die kleine InP-Substrate verwenden. Sie haben jedoch alle Vor- und Nachteile in Bezug auf Leistung, Kosten und heterogenes Integrationspotenzial. Das Imec hat die Aufgabe übernommen, die Vorteile und Herausforderungen für die verschiedenen Anwendungsfälle zu bewerten - sowohl für Infrastruktur als auch für mobile Geräte.

Ein erster Ansatz (Abbildung 4b) zur Herstellung von InP-on-Si-HBTs verwendet dehnungsreduzierte Pufferschichten, die direkt auf Si aufgebracht werden, um die achtprozentige Gitterfehlanpassung zwischen Si und InP auszugleichen. Anschließend wird das InP direkt auf diese Pufferschicht aufgewachsen. Die Möglichkeit, größere Wafergrößen zu verwenden, insbesondere in Fällen, in denen ein Teil des Si wiederverwendet werden kann, bietet einen erheblichen Kostenvorteil. Es sind jedoch Optimierungen erforderlich, um

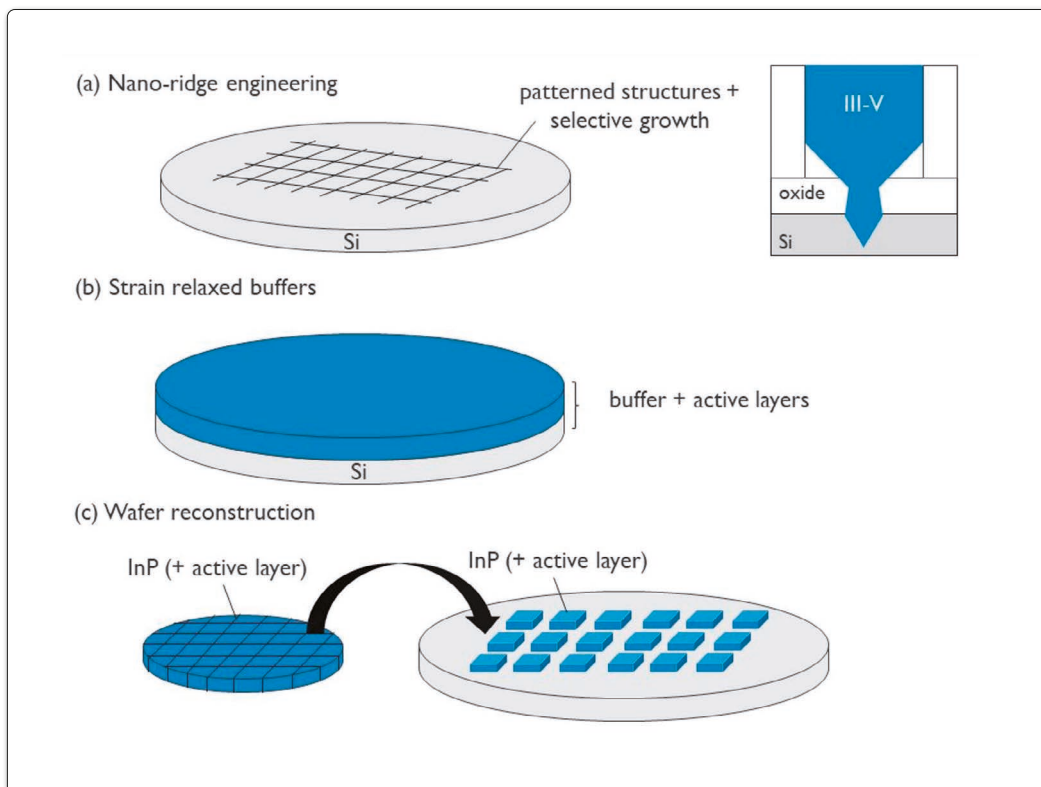


Bild 4: Schematische Darstellung der verschiedenen InP-on-Si-Wachstumsansätze: (a) Nano-Ridge-Engineering; (b) Blanket-Wachstum mit spannungsarmen Puffern und (c) Wafer-Rekonstruktion

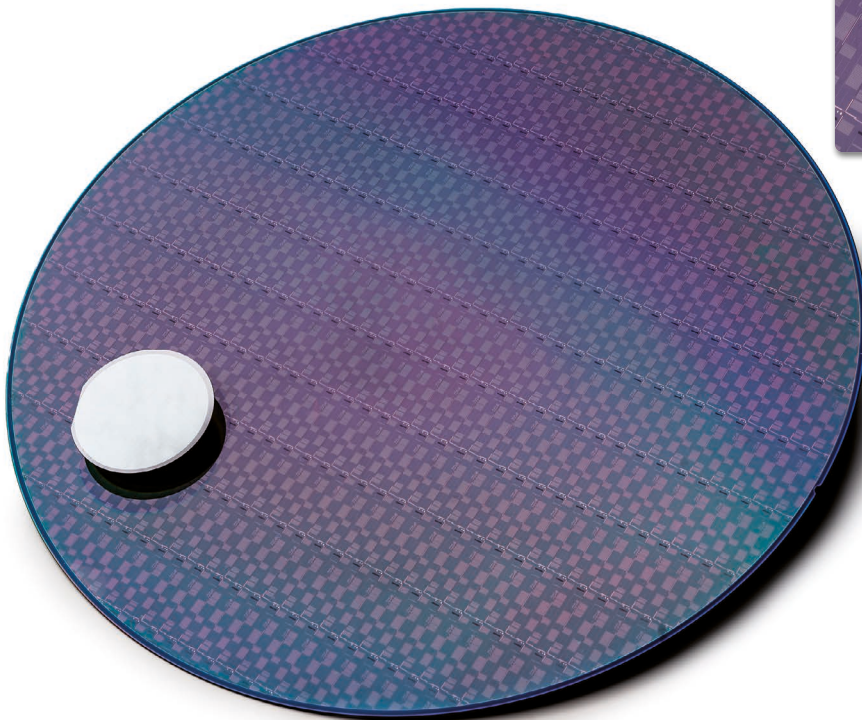


Bild 5: Ein Zwei-Zoll-InP-Wafer und ein 300-mm-Si-Wafer mit einem InP-NRE-HBT-Stapel

die Anzahl der Defekte weiter zu verringern.

Abweichend von diesem „pauschalen“ Wachstumsansatz schlägt imec das Nano-Ridge-Engineering (NRE) als alternative Technologie vor, um Defekte effizienter zu bewältigen (Abbildung 4a). NRE basiert auf dem selektiven Wachstum des III/V-Materials in vorstrukturierten Gräben in Si. Diese Gräben mit hohem Aspektverhältnis sind sehr effektiv, um die Defekte im schmalen unteren Teil einzufangen und das Wachstum von hochwertigem Material mit geringer Defektivität aus dem Graben heraus zu ermöglichen. Durch Überwachsen des Nanorückens wird dieser nach oben hin breiter und bildet eine solide Basis für einen Bauelementestapel. Die ersten Erkenntnisse aus einer GaAs/InGaP-Fallstudie werden die

Optimierung der angestrebten InGaAs/InP-NRE-HBT-Bauelemente leiten.

Neben dem direkten Wachstum kann InP auch mittels einer Wafer-Rekonstruktionstechnologie auf Si aufgebracht werden (Abbildung 4c). In diesem Fall werden hochwertige InP-Substrate – mit oder ohne die aktiven Schichten – während der Waferkonstitution in Kacheln zerlegt. Die Kacheln werden anschließend mit einem Die-to-Wafer-Bonding-Verfahren auf einen Si-Wafer aufgebracht. Die zentralen Herausforderungen liegen in der effizienten Übertragung der Materialien und der Entfernung des InP-Substrats, wofür mehrere Techniken in Betracht gezogen werden.

Auf dem Weg zur heterogenen Integration

Letztlich müssen die III/V-on-Si-Leistungsverstärker mit CMOS-basierten Komponenten kombiniert werden, die z.B. die Kalibrierung und Steuerung übernehmen. Imec prüft verschiedene heterogene Integrationsoptionen und wägt deren Vor- und Nachteile für verschiedene Anwendungsfälle ab.

Die fortschrittliche Laminat-substrat-Technologie ist die gängigste Art, verschiedene HF-Komponenten in ein System-in-Package zu integrieren, und Optimierungen, um sie für höhere Frequenzen geeignet zu machen, sind im Gange.

Außerdem erforscht imec fortschrittlichere heterogene Integrationsoptionen, einschließlich 2,5D-Interposer- und 3D-Integrationstechnologien.

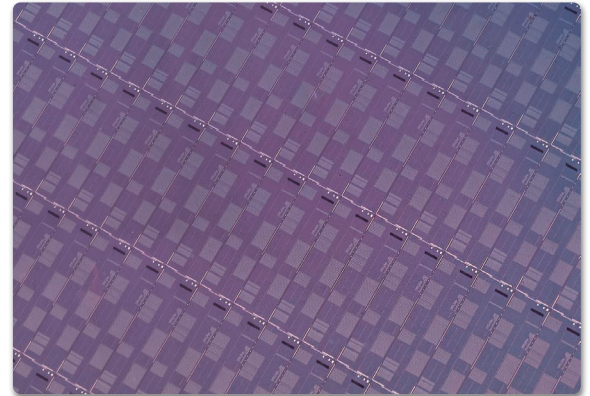


Bild 6: Vergrößerung des 300-mm-Si-Wafers, die die Dies mit den InP-NRE HBT-Strukturen zeigt

Insbesondere bei Frequenzen über 100 GHz ist zu beachten, dass das Antennenmodul die für den Transceiver zur Verfügung stehende Fläche zu definieren beginnt. Bei höheren Frequenzen nimmt die Wellenlänge ab, und die Fläche der Antennengruppe wird entsprechend größer. Oberhalb von 100 GHz wird die Größe der Antenne kleiner als die Größe des Front-End-Moduls, das mit zunehmender Frequenz kaum noch an Größe zunimmt. Eine interessante Option für große Antennengruppen-Konfigurationen ist die Verlegung des HF-Frontend-Moduls unter die Antennengruppe. Und hier kommen die 3D-Integrationstechnologien (entweder Die-to-Wafer oder Wafer-to-Wafer) ins Spiel, die kurze und gut definierte Verbindungen zwischen dem Frontend-Modul und den Antennenmodulen ermöglichen.

Das Wärmemanagement ist jedoch nach wie vor ein großes Problem bei der 3D-Integration, und die Fähigkeit, effektive Kühlkörper bereitzustellen, wird entscheidend sein. Heute führen wir bei imec eine umfassende STCO-Analyse (System-Technology-Co-Optimization) durch, um verschiedene Technologien für die 3D-Integration zu bewerten und die Technologieauswahl aus der Perspektive der Systemebene anzuleiten.

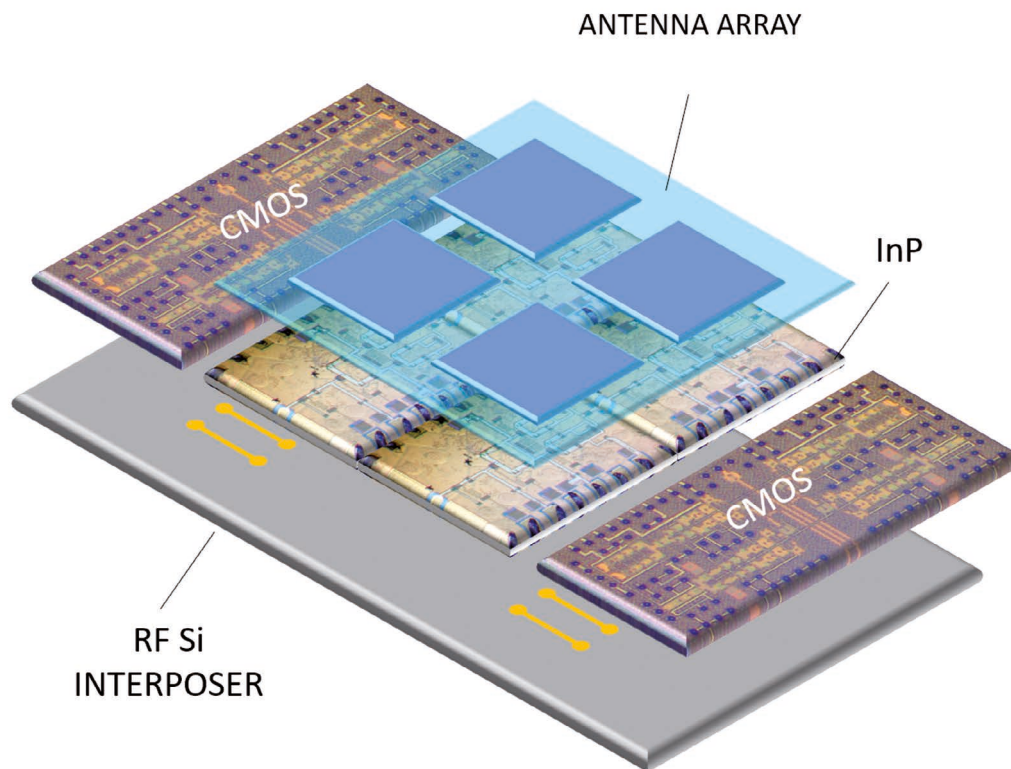


Bild 7: Schematische Darstellung eines RF-Si-Interposers mit integrierten InP- und CMOS-Bauelementen und einem Antennenarray in einem Gehäuse

Für Handheld-Geräte, bei denen eine geringere Anzahl von Antennen die Beschränkungen lockern kann, ist die 2,5D-Interposer-Technologie ein interessanter Ansatz. Bei dieser heterogenen Integrationsoption wird ein Schichtenstapel mit lithografisch definierten Verbindungen und sogar Durchkontaktierungen verwendet, um zwischen III/V- und CMOS-basierten Komponenten zu kommunizieren. In diesem Fall sitzen die III/V-Bauteile neben dem CMOS-Chip, was ein besseres Wärmemanagement ermöglicht, da beide Chips in direktem Kontakt mit einem Kühlkörper stehen können. Eine solche Architektur ermöglicht jedoch nur eine 1D-Strahlführung. Wir evaluieren derzeit Hardware-Implementierungen der 2,5D-Interposer-Technologie und untersuchen die optimalen Kombinationen von Substraten, Dielektrika und Umverteilungsschichten zur Minimierung von Verlusten. So haben wir beispielsweise eine erste Version einer RF-angepassten Si-Interposer-Technologie gezeigt, bei der ein Standard-Si-Substrat,

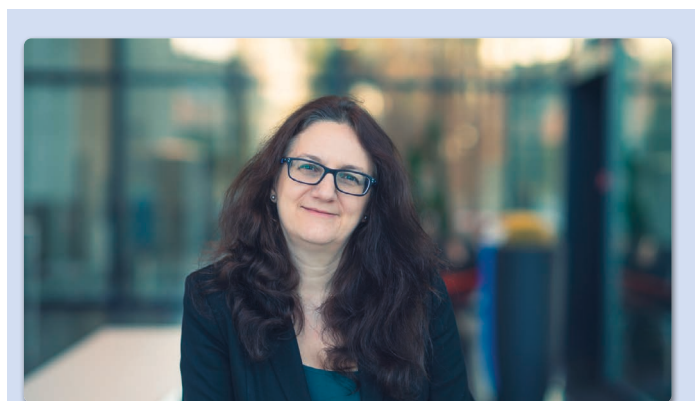
eine halbadditive Kupferverbindung und dicke, aufgeschleuderte Low-k-Dielektrika verwendet werden, die sehr geringe Verbindungsverluste aufweisen, sogar über 100 GHz.

Zusammenfassend lässt sich sagen, dass die jüngsten Bemühungen um eine Hochskalierung und Integration zeigen, dass GaN-on-Si und InP-on-Si realisierbare Technologien für drahtlose Kommunikationsanwendungen der nächsten Generation mit hoher Kapazität werden können.

Weitere Lektüre

Thermal modeling of GaN & InP RF devices with intrinsic account for nanoscale transport effects, B. Vermeersch et al., 2022 International Electron Devices Meeting (IEDM), DOI: 10.1109/IEDM45625.2022.10019370

III-V/III-N technologies for next-generation high-capacity wireless communication, N. Collaert et al., 2022 International Electron Devices Meeting, DOI: 10.1109/IEDM45625.2022.10019555 ◀



Die Autorin

Dr. Nadine Collaert ist Programmleiterin bei imec. Sie ist derzeit verantwortlich für das Advanced RF-Programm, das sich mit der heterogenen Integration von III-V/III-N-Bauelementen mit fortschrittlichen CMOS-Bauelementen befasst, um die Herausforderungen der nächsten Generation mobiler Kommunikation zu bewältigen. Zuvor war sie Programmleiterin des Programms „Logic Beyond Si“, das sich auf die Erforschung neuartiger CMOS-Bauelemente und Bauelemente und Systemansätze mit neuen

Materialien zur Erhöhung der Funktionalität konzentriert. Sie beschäftigt sich mit der Theorie, dem Design und der Technologie von FinFET-Bauelementen, neuartigen Speichern, Wandlern für biomedizinische Anwendungen sowie der Integration und Charakterisierung biokompatibler Materialien. Sie hat an der KU Leuven in Elektrotechnik promoviert, ist (Mit-)Autorin von mehr als 400 Veröffentlichungen und hält mehr als zehn Patente in den Bereichen Bauelementedesign und Prozesstechnologie.