

US Patent & Trademark Office

Patent Public Search | Text View

United States Patent Application Publication

20250266290

Kind Code

A1

Publication Date

August 21, 2025

Inventor(s)

SHEN; Shu-Wen et al.

Multigate Device Structure with Engineered Cladding and Method Making the Same

Abstract

The present disclosure provides a method of making a semiconductor device. The method includes forming a semiconductor stack on a substrate, wherein the semiconductor stack includes first semiconductor layers of a first semiconductor material and second semiconductor layers of a second semiconductor material alternatively stacked on the substrate; patterning the semiconductor stack and the substrate to form a trench and an active region being adjacent the trench; epitaxially growing a liner of the first semiconductor material on sidewalls of the trench and sidewalls of the active region; forming an isolation feature in the trench; performing a rapid thermal nitridation process, thereby converting the liner into a silicon nitride layer; and forming a cladding layer of the second semiconductor material over the silicon nitride layer.

Inventors: SHEN; Shu-Wen (Hsinchu City, TW), PENG; Yuan-Ching (Hsinchu, TW), CHEN; Chien-Hung (Hsinchu City, TW), LIN; Jheng-Wei (Hsinchu City, TW), YANG; Ji-Xuan (Hsinchu, TW), KUO; Jiun-Ming (Taipei City, TW)

Applicant: Taiwan Semiconductor Manufacturing Company, Ltd. (Hsinchu, TW)

Family ID: 1000008574725

Appl. No.: 19/067103

Filed: February 28, 2025

Related U.S. Application Data

parent US continuation 17805715 20220607 parent-grant-document US 12249539 child US 19067103

us-provisional-application US 63286306 20211206

Publication Classification

Int. Cl.: H01L21/762 (20060101); **H01L21/02** (20060101); **H10D30/01** (20250101); **H10D30/43** (20250101); **H10D30/67** (20250101); **H10D62/10** (20250101); **H10D64/01** (20250101)

U.S. Cl.:

CPC H01L21/76224 (20130101); **H01L21/02164** (20130101); **H01L21/0217** (20130101); **H01L21/02247** (20130101); **H01L21/02532** (20130101); **H01L21/02603** (20130101); **H10D30/014** (20250101); **H10D30/031** (20250101); **H10D30/43** (20250101); **H10D30/6735** (20250101); **H10D30/6757** (20250101); **H10D62/121** (20250101); **H10D64/017** (20250101);

Background/Summary

[0001] This application is a continuation of U.S. patent application Ser. No. 17/805,715, filed Jun. 7, 2022, which claims priority to U.S. Provisional Patent Application Ser. No. 63/286,306 filed Dec. 6, 2021, the entire disclosures of which are hereby incorporated herein by reference.

BACKGROUND

[0002] The electronics industry has experienced an ever-increasing demand for smaller and faster electronic devices that are simultaneously able to support a greater number of increasingly complex and sophisticated functions. To meet these demands, there is a continuing trend in the integrated circuit (IC) industry to manufacture low-cost, high-performance, and low-power ICs. Thus far, these goals have been achieved in large part by reducing IC dimensions (for example, minimum IC feature size), thereby improving production efficiency and lowering associated costs. However, such scaling has also increased complexity of the IC manufacturing processes. Thus, realizing continued advances in IC devices and their performance requires similar advances in IC manufacturing processes and technology.

[0003] Recently, multigate devices have been introduced to improve gate control. Multigate devices have been observed to increase gate-channel coupling, reduce OFF-state current, and/or reduce short-channel effects (SCEs). One such multigate device is the gate-all around (GAA) device, which includes a gate structure that can extend, partially or fully, around a channel region to provide access to the channel region on at least two sides. GAA devices enable aggressive scaling down of IC technologies, maintaining gate control and mitigating SCEs, while seamlessly integrating with conventional IC manufacturing processes. As GAA devices continue to scale, challenges have arisen when fabricating a gate structure for a GAA device, which challenges have been observed to degrade GAA device performance and increase GAA processing complexity. For example, a cladding layer formed around an active region may have footing issue, which eventually cause residual defects during gate replacement. In another example, the cladding layer may introduce diffusion into channels, which reduces carrier mobility and degrade the device performance. Accordingly, although existing GAA devices and methods for fabricating such have been generally adequate for their intended purposes, they have not been entirely satisfactory in all respects.

Description

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] The present disclosure is best understood from the following detailed description when read with the accompanying figures. It is emphasized that, in accordance with the standard practice in the industry, various features are not drawn to scale and are used for illustration purposes only. In

fact, the dimensions of the various features may be arbitrarily increased or reduced for clarity of discussion.

[0005] FIG. 1A is a flowchart of a method for fabricating a multigate device according to various aspects of the present disclosure.

[0006] FIGS. 1B, 1C, 1D and 1E are flowcharts of a method for fabricating a multigate device according to various embodiments of the present disclosure.

[0007] FIGS. 2A-4A, FIGS. 2B-4B, FIGS. 2C-4C, and FIGS. 2D-4D are fragmentary diagrammatic views of a multigate device, in portion or entirety, at various fabrication stages (such as those associated with the method in FIGS. 1A-1E) according to various aspects of the present disclosure.

[0008] FIGS. 5A-5F are fragmentary sectional views of the multigate device, in portion or entirety, constructed according to some embodiments.

[0009] FIGS. 6A-6H are fragmentary sectional views of the multigate device, in portion or entirety, constructed according to some embodiments.

[0010] FIGS. 7A-7F are fragmentary sectional views of the multigate device, in portion or entirety, constructed according to some embodiments.

[0011] FIG. 7G is a diagrammatic view of a chemical structure constructed according to some embodiments.

[0012] FIGS. 8A-8D are fragmentary sectional views of the multigate device, in portion or entirety, constructed according to some embodiments.

[0013] FIGS. 9A, 10A, 11A, 12A, 13A, 14A, 15A, 16A, 17A and 18A, FIGS. 9B, 10B, 11B, 12B, 13B, 14B, 15B, 16B, 17B and 18B, FIGS. 9C, 10C, 11C, 12C, 13C, 14C, 15C, 16C, 17C and 18C, and FIGS. 9D, 10D, 11D, 12D, 13D, 14D, 15D, 16D, 17D and 18D are fragmentary diagrammatic views of a multigate device, in portion or entirety, at various fabrication stages (such as those associated with the method in FIGS. 1A-1E) according to various aspects of the present disclosure.

[0014] FIGS. 9E and 9F are fragmentary diagrammatic views of the multigate device, in portion or entirety according to various aspects of the present disclosure.

[0015] FIG. 16E is a fragmentary diagrammatic view of the multigate device, in portion or entirety according to various aspects of the present disclosure.

[0016] FIG. 17E is a fragmentary diagrammatic view of the multigate device, in portion or entirety according to various aspects of the present disclosure.

[0017] FIG. 18E is a fragmentary diagrammatic view of the multigate device, in portion or entirety according to various aspects of the present disclosure.

DETAILED DESCRIPTION

[0018] The present disclosure relates generally to integrated circuit devices, and more particularly, to multigate devices, such as gate-all-around (GAA) devices.

[0019] The following disclosure provides many different embodiments, or examples, for implementing different features. Reference numerals and/or letters may be repeated in the various examples described herein. This repetition is for the purpose of simplicity and clarity and does not in itself dictate a relationship between the various disclosed embodiments and/or configurations. Further, specific examples of components and arrangements are described below to simplify the present disclosure. These are, of course, merely examples and are not intended to be limiting. For example, the formation of a first feature over or on a second feature in the description that follows may include embodiments in which the first and second features are formed in direct contact and may also include embodiments in which additional features may be formed between the first and second features, such that the first and second features may not be in direct contact. Moreover, the formation of a feature on, connected to, and/or coupled to another feature in the present disclosure may include embodiments in which the features are formed in direct contact, and may also include embodiments in which additional features may be formed interposing the features, such that the features may not be in direct contact.

[0020] In addition, the present disclosure may repeat reference numerals and/or letters in the various examples. This repetition is for the purpose of simplicity and clarity and does not in itself dictate a relationship between the various embodiments and/or configurations discussed. Moreover, the formation of a feature on, connected to, and/or coupled to another feature in the present disclosure that follows may include embodiments in which the features are formed in direct contact, and may also include embodiments in which additional features may be formed interposing the features, such that the features may not be in direct contact. In addition, spatially relative terms, for example, “lower,” “upper,” “horizontal,” “vertical,” “above,” “over,” “below,” “beneath,” “up,” “down,” “top,” “bottom,” etc. as well as derivatives thereof (e.g., “horizontally,” “downwardly,” “upwardly,” etc.) are used for ease of the present disclosure of one feature relationship to another feature. The spatially relative terms are intended to cover different orientations of the device including the features. Still further, when a number or a range of numbers is described with “about,” “approximate,” and the like, the term is intended to encompass numbers that are within a reasonable range including the number described, such as within $\pm 10\%$ of the number described, or other values as understood by person skilled in the art. For example, the term “about 5 nm” encompasses the dimension range from 4.5 nm to 5.5 nm.

[0021] The disclosed FET structure is formed on a planar active region as a planar FET device, and alternatively is formed on a three-dimensional (3D) structure, such as multi-gate FET devices. Examples of multi-gate devices include fin-like field effect transistors (FinFETs) having fin-like structures and multi-bridge-channel (MBC). An MBC transistor has a gate structure that can extend, partially or fully, around a channel region to provide access to the channel region on two or more sides. Because its gate structure surrounds the channel regions, an MBC transistor may also be referred to as a surrounding gate transistor (SGT) or a gate-all-around (GAA) transistor having a plurality of channel members vertically stacked.

[0022] FIG. 1A illustrates a flowchart of a method **100** for fabricating a multi-gate device according to various aspects of the present disclosure. The method **100** begins with a workpiece having a semiconductor substrate. In some embodiments, method **100** fabricates a multi-gate device, such as a multi-gate device that includes p-type GAA transistors and n-type GAA transistors. In some embodiments, method **100** fabricates a multi-gate device that includes first GAA transistors and second GAA transistors with different characteristics, such as different functions (e.g., logic device or memory device) or different conductivity type (e.g., n-type transistor or p-type transistor). In the disclosed structure and the method making the same, the device structure includes an engineered cladding layer designed differently to optimize respective device performance, including reduced diffusion, and reduced cladding defects, such as footing issue. Particularly, the method includes forming a diffusion blocking layer before forming a cladding layer according to various embodiments. In some embodiments, the cladding layer is epitaxially grown to selectively deposit on the surfaces of the active region relative to the surfaces of an isolation feature, such as shallow trench isolation (STI) feature.

[0023] In some embodiments, method **100** fabricates a multi-gate device that includes p-type GAA transistors and n-type GAA transistors. At block **101**, a semiconductor layer stack is formed over a substrate. The semiconductor layer stack includes first semiconductor layers and second semiconductor layers stacked vertically in an alternating configuration. In some embodiments, the operation **101** includes depositing various semiconductor materials (such as alternatively silicon and silicon germanium) by epitaxial growth. In some embodiments, doped wells, such as n-type doped wells and p-type doped wells are formed by ion implantations before the formation of the semiconductor stack. At block **102**, the semiconductor layer stack is patterned to form a trench and an active region (also referred to as semiconductor fin) being surrounded by the trench. A hard mask may be formed during the process to pattern the semiconductor layer stack. At block **103**, an engineered cladding layer is formed on sidewalls of the active region. The cladding layer provides paths to selectively remove the second semiconductor layers and release the channels. The block

103 includes multiple operations to form the engineered cladding layer according to various embodiments, which are to be further described at later stage.

[0024] In some embodiments, dielectric fins may be formed on the substrate among the fins. Dielectric fins have similar profile as semiconductor fins but consist dielectric material(s) with benefits, such as tuning the fin density. At block **104**, a gate structure is formed over the semiconductor layer stack. The gate structure includes a dummy gate stack and gate spacers. A lightly doped drain (LDD) implantation may be implemented, and the cladding layer may be anisotropically etched to remove the portions disposed on the isolation features, such as shallow trench isolation (STI) features. The anisotropic etch may include plasma etch. At block **106**, portions of the semiconductor layer stack in source/drain regions are removed to form source/drain recesses. At block **108**, inner spacers are formed along sidewalls of the second semiconductor layers in the semiconductor layer stack. At block **110**, epitaxial source/drain (S/D) features are formed in the source/drain recesses. At block **112**, an interlayer dielectric (ILD) layer is formed over the epitaxial source/drain features. At block **114**, the dummy gate stack is removed, thereby forming a gate trench that expose the semiconductor layer stack. At block **116**, the cladding layer and the first semiconductor layers are selectively removed from the semiconductor layer stack exposed in the gate trench, thereby forming gaps between the second semiconductor layers. At block **120**, a gate stack is formed in the gate trench and is wrapping around each of the first semiconductor layers. At block **122**, other fabrication processes, including forming an interconnect structure, are performed on the workpiece. Additional steps can be provided before, during, and after method **100**, and some of the steps described can be moved, replaced, or eliminated for additional embodiments of method **100**. The discussion that follows illustrates various embodiments of nanowire-based (or nanostructure-based) integrated circuit devices that can be fabricated according to method **100**.

[0025] FIG. **1B** illustrates a flowchart of a method **103** for fabricating an engineered cladding layer according to some embodiments of the present disclosure. The method **103** includes forming the cladding layer and an isolation feature in the trench. At block **132**, a liner of a semiconductor material is formed on sidewalls of the semiconductor fin. The liner is a semiconductor layer, such as a silicon layer, selectively grown on semiconductor surfaces. In the disclosed embodiment, the liner is epitaxially grown on sidewalls of the semiconductor fin. At block **134**, the isolation feature is formed in the trench. The isolation feature includes one or more dielectric material and provides isolation of semiconductor fin from other active regions (such as other semiconductor fins). In the disclosed embodiment, the isolation feature is a shallow trench isolation (STI) feature and is formed by deposition, chemical mechanical polishing (CMP) and etch. In furtherance of the embodiment, the process to form the STI feature includes depositing one or more dielectric material to fill the trench; performing a CMP process to remove excessive dielectric material and planarize the top surface; and selective etching to recess the dielectric material so that the active region is extruded above the top surface of the STI feature. Therefore, the active region is also referred to as semiconductor fin. At block **136**, a dielectric layer is formed over the sidewalls of the semiconductor fin, particularly on the liner. In the disclosed embodiment, the dielectric layer is a silicon oxide layer formed by a suitable technique, such as thermal oxidation. At block **138**, a rapid thermal nitridation (RTN) process is applied to the workpiece, thereby partially or completely converting the silicon liner into a silicon nitride layer. The silicon nitride layer functions as a diffusion blocking layer to prevent the diffusion of the subsequent formed cladding layer into the semiconductor layer stack. At block **140**, a cladding layer of a semiconductor material, such as silicon germanium, is formed on the silicon nitride layer and the dielectric layer. the cladding layer and the liner are different in composition. In this case, the cladding layer is a semiconductor layer in an amorphous structure due to the dielectric surfaces. At block **142**, an anisotropic etching process, such as plasma etching, is applied to the cladding layer to substantially remove the portions of the cladding layer disposed on the STI feature.

[0026] FIG. 1C illustrates a flowchart of a method **103** for fabricating an engineered cladding layer according to some embodiments of the present disclosure. The method **103** in FIG. 1C is similar to the method **103** in FIG. 1B except for the operation **146**. The similar descriptions are not repeated here for simplicity. After the block **138**, the block **146** is implemented. At block **146**, the dielectric layer (such as silicon oxide layer) is selectively removed by an etching, such as a wet etching, a dry etching or a combination thereof. Thus, the silicon nitride layer is exposed. Therefore, the subsequent formed cladding layer includes first portions disposed on the liner and second portions disposed on the silicon nitride layer. The first portions of the cladding layer are in an amorphous structure and the second portions of the cladding layer are in a crystalline structure.

[0027] FIG. 1D illustrates a flowchart of a method **103** for fabricating an engineered cladding layer according to some embodiments of the present disclosure. The method **103** in FIG. 1D is similar to the method **103** in FIG. 1B except for the operation **148**. The similar descriptions are not repeated here for simplicity. The block **148** is implemented after the block **136**. At block **148**, an in-situ carbon (ISC) process is performed to form a carbon-rich layer on the dielectric layer and the liner. The carbon-rich layer includes nitrogen and is carbon-rich nitride layer or carbon-rich dielectric layer. The carbon-rich layer functions as a diffusion blocking layer to prevent the diffusion of the subsequent formed cladding layer into the semiconductor layer stack. Therefore, the subsequent formed cladding layer is in an amorphous structure due to the characteristics of the surfaces. At block **142**, an anisotropic etching process, such as plasma etching, is applied to the cladding layer to substantially remove the portions of the cladding layer disposed on the STI feature. In some embodiments, the etching process is applied with 100% over etch so that the portions of cladding layer disposed on the STI feature are substantially removed. In some embodiments, the etching process is applied with less than 100% over etch (such as 50% over etch) so that the portions of cladding layer disposed on the STI feature are remained. In this case, the portions of the cladding layer disposed on the carbon-rich layer and the STI feature are in amorphous structure and the portion disposed on the liner in the corner is in a crystalline structure.

[0028] FIG. 1E illustrates a flowchart of a method **103** for fabricating an engineered cladding layer according to some embodiments of the present disclosure. The method **103** includes forming an isolation feature in the trench and forming a cladding layer. At block **132**, a liner of a semiconductor material is formed on sidewalls of the semiconductor fin. The liner is a semiconductor layer, such as a silicon layer, selectively grown on semiconductor surfaces. In the disclosed embodiment, the liner is epitaxially grown on sidewalls of the semiconductor fin. At block **134**, the isolation feature is formed in the trench. The isolation feature includes one or more dielectric material and provides isolation of semiconductor fin from other active regions (such as other semiconductor fins). In the disclosed embodiment, the isolation feature is a shallow trench isolation (STI) feature and is formed by deposition, chemical mechanical polishing (CMP) and etch. In furtherance of the embodiment, the process to form the STI feature includes depositing one or more dielectric material to fill the trench; performing a CMP process to remove excessive dielectric material and planarize the top surface; and selective etching to recess the dielectric material so that the active region is extruded above the top surface of the STI feature. Therefore, the active region is also referred to as semiconductor fin. At block **150**, an etching process is performed to remove the liner and oxide materials on the sidewalls of the semiconductor fins. At block **152**, a cladding layer of a semiconductor material, such as silicon germanium, is selectively formed on the sidewalls of the semiconductor fins. The cladding layer is formed by selective epitaxial growth. In this case, the cladding layer is a semiconductor layer in a crystalline structure. Furthermore, the cladding layer is selectively disposed on the sidewalls of the semiconductor fins and is free from the surface of the isolation features, such as STI features. Therefore, the anisotropic etching process applied to the cladding layer is skipped and the footing issues of the cladding layer are eliminated.

[0029] FIGS. 2A-18E are fragmentary diagrammatic views of a multigate device (or a workpiece)

200, in portion or entirety, at various fabrication stages (such as those associated with method **100** in FIG. **1A**) according to various aspects of the present disclosure. In particular, FIGS. **2A-4A** and **9A-18A** are top views of multigate device **200** in an X-Y plane; FIGS. **2B-4B** and **9B-18B** are diagrammatic cross-sectional views of multigate device **200** in an X-Z plane along lines B-B' respectively of FIGS. **2A-4A** and **9A-18A**, FIGS. **2C-4C** and **9C-18C** are diagrammatic cross-sectional views of multigate device **200** in a Y-Z plane along lines C-C' respectively of FIGS. **2A-4A** and **9A-18A**; and FIGS. **2D-4D** and **9D-18D** are diagrammatic cross-sectional views of multigate device **200** in the Y-Z plane along lines D-D' respectively of FIGS. **2A-4A** and **9A-18A**. [0030] FIGS. **5A-5F** are fragmentary sectional views of the multigate device **200**, in portion or entirety, constructed according to some embodiments. FIGS. **6A-6H** are fragmentary sectional views of the multigate device, in portion or entirety, constructed according to some embodiments. FIGS. **7A-7F** are fragmentary sectional views of the multigate device, in portion or entirety, constructed according to some embodiments. FIG. **7G** is a diagrammatic view of a chemical structure constructed according to some embodiments. FIGS. **8A-8D** are fragmentary sectional views of the multigate device, in portion or entirety, constructed according to some embodiments. [0031] FIG. **9E**, and FIG. **9F** are fragmentary diagrammatic views of the multigate device **200**, in portion or entirety according to various aspects of the present disclosure. In particular, each of FIG. **9E** is a diagrammatic cross-sectional view of multigate device **200** in the X-Z plane along lines B-B' of FIG. **9A** constructed according to various embodiments. FIG. **9F** is a diagrammatic cross-sectional view of multigate device **200** in the X-Z plane along lines B-B' of FIG. **9A** constructed according to various embodiments.

[0032] FIG. **16E** is a fragmentary diagrammatic view of the multigate device **200** in the X-Z plane along lines B-B' of FIG. **16A**; in portion or entirety according to various aspects of the present disclosure. FIG. **17E** is a diagrammatic cross-sectional view of multigate device **200** in the X-Z plane along lines B-B' of FIG. **17A** constructed according to some embodiments. FIG. **18E** is a diagrammatic cross-sectional view of multigate device **200** in the X-Z plane along lines B-B' of FIG. **18A** constructed according to some embodiments.

[0033] Multigate device **200** may be included in a microprocessor, a memory, and/or other IC device. In some embodiments, multigate device **200** is a portion of an IC chip, a system on chip (SoC), or portion thereof, that includes various passive and active microelectronic devices such as resistors, capacitors, inductors, diodes, p-type field effect transistors (PFETs), n-type field effect transistors (NFETs), metal-oxide semiconductor field effect transistors (MOSFETs), complementary metal-oxide semiconductor (CMOS) transistors, bipolar junction transistors (BJTs), laterally diffused MOS (LDMOS) transistors, high voltage transistors, high frequency transistors, other suitable components, or combinations thereof. In some embodiments, multigate device **200** is included in a non-volatile memory, such as a non-volatile random-access memory (NVRAM), a static random-access memory (SRAM), a flash memory, an electrically erasable programmable read only memory (EEPROM), an electrically programmable read-only memory (EPROM), other suitable memory type, or combinations thereof. Various figures have been simplified for the sake of clarity to better understand the inventive concepts of the present disclosure. Additional features can be added in multigate device **200**, and some of the features described below can be replaced, modified, or eliminated in other embodiments of multigate device **200**. Particularly, the substrate **202** includes a first region for a logic circuit (also referred to as a logic circuit region) and a second region for a SRAM circuit (also referred to as a SRAM circuit region).

[0034] Turning to FIGS. **2A-2D**, multigate device **200** includes a substrate (e.g., wafer) **202**. In the depicted embodiment, substrate **202** includes silicon. Alternatively, or additionally, substrate **202** includes another elementary semiconductor, such as germanium; a compound semiconductor, such as silicon carbide, gallium arsenide, gallium phosphide, indium phosphide, indium arsenide, and/or indium antimonide; an alloy semiconductor, such as silicon germanium (SiGe), GaAsP, AlInAs, AlGaAs, GaInAs, GaInP, and/or GaInAsP; or combinations thereof. Alternatively, substrate **202** is

a semiconductor-on-insulator substrate, such as a silicon-on-insulator (SOI) substrate, a silicon germanium-on-insulator (SGOI) substrate, or a germanium-on-insulator (GOI) substrate. Semiconductor-on-insulator substrates can be fabricated using separation by implantation of oxygen (SIMOX), wafer bonding, and/or other suitable methods. Substrate **202** can include various doped regions depending on design requirements of multigate device **200**. In the depicted embodiment, substrate **202** includes a p-type doped region (referred to hereinafter as a p-well), which can be configured for n-type GAA transistors, and an n-type doped region (referred to hereinafter as an n-well), which can be configured for p-type GAA transistors. N-type doped regions, such as n-well, are doped with n-type dopants, such as phosphorus, arsenic, other n-type dopant, or combinations thereof. P-type doped regions, such as p-well, are doped with p-type dopants, such as boron, indium, other p-type dopant, or combinations thereof. In some implementations, substrate **202** includes doped regions formed with a combination of p-type dopants and n-type dopants. The various doped regions can be formed directly on and/or in substrate **202**, for example, providing a p-well structure, an n-well structure, a dual-well structure, a raised structure, or combinations thereof. An ion implantation process, a diffusion process, and/or other suitable doping process can be performed to form the various doped regions.

[0035] A semiconductor layer stack **205** is formed over substrate **202**, where semiconductor layer stack **205** includes first semiconductor layers **210** and second semiconductor layers **215** stacked vertically (e.g., along the z-direction) in an interleaving or alternating configuration from a surface of substrate **202**. In some embodiments, semiconductor layers **210** and semiconductor layers **215** are epitaxially grown in the depicted interleaving and alternating configuration. For example, a first one of semiconductor layers **210** is epitaxially grown on substrate, a first one of semiconductor layers **215** is epitaxially grown on the first one of semiconductor layers **210**, a second one of semiconductor layers **210** is epitaxially grown on the first one of semiconductor layers **215**, and so on until semiconductor layers stack **205** has a desired number of first semiconductor layers **210** and second semiconductor layers **215**. In such embodiments, first semiconductor layers **210** and second semiconductor layers **215** can be referred to as epitaxial layers. In some embodiments, epitaxial growth of first semiconductor layers **210** and second semiconductor layers **215** is achieved by a molecular beam epitaxy (MBE) process, a chemical vapor deposition (CVD) process, a metalorganic chemical vapor deposition (MOCVD) process, other suitable epitaxial growth process, or combinations thereof.

[0036] A composition of first semiconductor layers **210** is different than a composition of second semiconductor layers **215** to achieve etching selectivity and/or different oxidation rates during subsequent processing. In some embodiments, first semiconductor layers **210** have a first etch rate to an etchant and second semiconductor layers **215** have a second etch rate to the etchant, where the second etch rate is less than the first etch rate. In some embodiments, first semiconductor layers **210** have a first oxidation rate and second semiconductor layers **215** have a second oxidation rate, where the second oxidation rate is less than the first oxidation rate. In the depicted embodiment, first semiconductor layers **210** and second semiconductor layers **215** include different materials, constituent atomic percentages, constituent weight percentages, thicknesses, and/or characteristics to achieve desired etching selectivity during an etching process, such as an etching process implemented to form suspended channel layers in channel regions of multigate device **200**. For example, where first semiconductor layers **210** include silicon germanium and second semiconductor layers **215** include silicon, a silicon etch rate of semiconductor layers **215** is less than a silicon germanium etch rate of semiconductor layers **210** by the etchant used in the later channel-releasing process. In some embodiments, semiconductor layers **210** and semiconductor layers **215** can include the same material but with different constituent atomic percentages to achieve the etching selectivity and/or different oxidation rates. For example, semiconductor layers **210** and semiconductor layers **215** can include silicon germanium, where semiconductor layers **210** have a first silicon atomic percent and/or a first germanium atomic percent and semiconductor

layers **215** have a second, different silicon atomic percent and/or a second, different germanium atomic percent. The present disclosure contemplates that semiconductor layers **210** and semiconductor layers **215** include any combination of semiconductor materials that can provide desired etching selectivity, desired oxidation rate differences, and/or desired performance characteristics (e.g., materials that maximize current flow), including any of the semiconductor materials disclosed herein.

[0037] As described further below, semiconductor layers **215** or portions thereof form channel regions of multigate device **200**. In the depicted embodiment, semiconductor layer stack **205** includes four semiconductor layers **210** and four semiconductor layers **215** configured to form four semiconductor layer pairs disposed over substrate **202**, each semiconductor layer pair having a respective first semiconductor layer **210** and a respective second semiconductor layer **215**. After undergoing subsequent processing, such configuration will result in multigate device **200** having four channels. However, the present disclosure contemplates embodiments where semiconductor layer stack **205** includes more or less semiconductor layers, for example, depending on a number of channels desired for multigate device **200** (e.g., a GAA transistor) and/or design requirements of multigate device **200**. For example, semiconductor layer stack **205** can include two to ten semiconductor layers **210** and two to ten semiconductor layers **215**. In furtherance of the depicted embodiment, semiconductor layers **210** have a thickness t_1 and semiconductor layers **215** have a thickness t_2 , where thickness t_1 and thickness t_2 are chosen based on fabrication and/or device performance considerations for multigate device **200**. For example, thickness t_1 can be configured to define a desired distance (or gap) between adjacent channels of multigate device **200** (e.g., between semiconductor layers **215**), thickness t_2 can be configured to achieve desired thickness of channels of multigate device **200**, and both thickness t_1 and thickness t_2 can be configured to achieve desired performance of multigate device **200**. In some embodiments, thickness t_1 and thickness t_2 are about 1 nm to about 10 nm.

[0038] Turning to FIGS. 3A-3D, semiconductor layer stack **205** is patterned to form semiconductor fins **218** and trenches **219** surrounding various semiconductor fins with isolation features to be formed therein for isolation. Semiconductor fins **218** function as active regions for various IC devices, such as transistors, to be formed therein. In some embodiments, semiconductor fins **218** further include semiconductor fins **218A** for n-type transistors in region **202A** and semiconductor fins **218B** for p-type transistors in region **202B**. Only one semiconductor fin **218A** and one semiconductor fin **218B** are illustrated in FIG. 3A. However, the present disclosure contemplates embodiments where semiconductor fins **218A** includes a plurality of semiconductor fins **218A** and semiconductor fins **218B** includes a plurality of semiconductor fins **218B**, for example, depending on a number of n-type GAA transistors and a number of p-type GAA transistors desired for multigate device **200** and/or design requirements of multigate device **200**. Fins **218A**, **218B** include a substrate portion (i.e., a portion of substrate **202**) and a semiconductor layer stack portion (i.e., a remaining portion of semiconductor layer stack **205** including semiconductor layers **210** and semiconductor layers **215**). Fins **218A**, **218B** extend substantially parallel to one another along a y-direction, having a length defined in the y-direction, a width defined in an x-direction, and a height defined in a z-direction. In some implementations, a lithography and/or etching process is performed to pattern semiconductor layer stack **205** to form fins **218A**, **218B**. The lithography process can include forming a resist layer over semiconductor layer stack **205** (for example, by spin coating), performing a pre-exposure baking process, performing an exposure process using a mask, performing a post-exposure baking process, and performing a developing process. During the exposure process, the resist layer is exposed to radiation energy (such as ultraviolet (UV) light, deep UV (DUV) light, or extreme UV (EUV) light), where the mask blocks, transmits, and/or reflects radiation to the resist layer depending on a mask pattern of the mask and/or mask type (for example, binary mask, phase shift mask, or EUV mask), such that an image is projected onto the resist layer that corresponds with the mask pattern. Since the resist layer is sensitive to radiation

energy, exposed portions of the resist layer chemically change, and exposed (or non-exposed) portions of the resist layer are dissolved during the developing process depending on characteristics of the resist layer and characteristics of a developing solution used in the developing process. After development, the patterned resist layer includes a resist pattern that corresponds with the mask. The etching process removes portions of semiconductor layer stack **205** using the patterned resist layer as an etch mask. In some embodiments, the patterned resist layer is formed over a hard mask layer disposed over semiconductor layer stack **205**, a first etching process removes portions of the hard mask layer to form a patterned hard mask layer, and a second etching process removes portions of semiconductor layer stack **205** using the patterned hard mask layer as an etch mask. The etching process can include a dry etching process, a wet etching process, other suitable etching process, or combinations thereof. In some embodiments, the etching process is a reactive ion etching (RIE) process. After the etching process, the patterned resist layer (and, in some embodiments, a hard mask layer) is removed, for example, by a resist stripping process or other suitable process. In the depicted embodiments, fins **218A**, **218B** are formed by a multiple patterning process, such as a double patterning lithography (DPL) process (for example, a lithography-etch-lithography-etch (LELE) process, a self-aligned double patterning (SADP) process, a spacer-is-dielectric (SID) SADP process, other double patterning process, or combinations thereof), a triple patterning process (for example, a lithography-etch-lithography-etch-lithography-etch (LELELE) process, a self-aligned triple patterning (SATP) process, other triple patterning process, or combinations thereof), other multiple patterning process (for example, self-aligned quadruple patterning (SAQP) process), or combinations thereof. In some embodiments, directed self-assembly (DSA) techniques are implemented while patterning semiconductor layer stack **205**. Further, in some embodiments, the exposure process can implement maskless lithography, electron-beam (e-beam) writing, and/or ion-beam writing for patterning the resist layer. In the depicted embodiment, the operation to form fins **218A** and **218B** includes at least one patterning procedure (or processes) that includes one lithography process and one etching process, resulting in the trenches **219**.

[0039] Turning to FIGS. **4A-4D**, an engineered cladding layer **220** is formed on sidewalls of the semiconductor fins **218** (such as **218A** and **218B**). Isolation features **230**, such as shallow trench isolation (STI) features, are formed in the trenches **219** and are surrounding semiconductor fins for isolation. Particularly, the engineered cladding layer **220** is configured and formed with characteristics for improved device performance and fabrication enhancement, such as reduced diffusion of the cladding layer into the channels and reduction of the cladding footing issues. The operations to form the engineered cladding layer **220** and isolation features **230** are further described below in detail according to various embodiments.

[0040] Some embodiments are described with reference to FIGS. **5A-5F**. The method may include operations described in FIG. **1B**. Turning to FIG. **5A**, a liner of a semiconductor material **402** is formed on sidewalls of the semiconductor fins **218** (such as **218A** and **218B**) and the trenches **219**. The liner **402** includes a semiconductor material, such as silicon. The liner **402** is formed by a suitable method, such as epitaxial growth. In some embodiments, the liner **402** includes a thickness ranging 1 nm and 2 nm. The liner **402** extends from the top surface of the semiconductor fins **218** to the bottom surface of the trenches **219**.

[0041] Turning to FIG. **5B**, isolation features **230**, such as STI features, are formed in the trenches **219**. In the disclosed embodiment, STI features are formed by deposition, CMP and etch. In furtherance of the embodiment, the process to form the STI features includes depositing one or more dielectric material (such as silicon oxide, silicon nitride, low-k dielectric material or a combination thereof) to fill the trench; performing a CMP process to remove excessive dielectric material and planarize the top surface; and selective etching to recess the dielectric material so that the active region is extruded above the top surface of the STI feature.

[0042] Turning to FIG. **5C**, a dielectric layer **404**, such as silicon oxide layer, is formed over the sidewalls of the semiconductor fins **218** in the trenches **219**. In the disclosed embodiment, the

dielectric layer **404** is formed on the liner **402**. The dielectric layer **404** has a thickness ranging between 0.8 nm and 1.5 nm. The dielectric layer **404** is formed by any suitable method, such as thermal oxidation, CVD, atomic layer deposition (ALD), or a combination thereof.

[0043] Turning to FIG. 5D, a rapid thermal nitridation (RTN) process is applied to the workpiece **200**, thereby partially or completely converting the liner **402** of silicon into a silicon nitride layer **406**. The silicon nitride layer **406** functions as a diffusion blocking layer to prevent the diffusion of the subsequent formed cladding layer into the semiconductor layer stack **205**. The RTN process is implemented at an elevated temperature with a precursor having nitrogen-containing chemical, such as N.sub.2, ammonia (NH₃), other suitable gas or a combination thereof. The elevated temperature may range between 700° C. and 900° C. The silicon nitride layer may have a thickness ranging between 5 Angstrom and 20 Angstrom. In some embodiments, the partial pressure of the nitrogen-containing chemical ranges between 5 torr and 500 torr during the RTN process and the RTN process duration ranges between 30 sec and 60 sec.

[0044] Turning to FIG. 5E, a cladding layer **408** of a semiconductor material, such as silicon germanium, is formed on the silicon nitride layer **406** and the dielectric layer **404**. The cladding layer **408** and the liner **402** are different in composition. The cladding layer **408** may include a semiconductor material similar to that of the first semiconductor layers **210** in composition. In the depicted embodiment, the cladding layer **408** includes silicon germanium. The cladding layer **408** provides paths to etch the first semiconductor layers **210** and is removed with the first semiconductor layers **210** during a channel-release process at later stage. In the present embodiment, the cladding layer **408** is a semiconductor layer in an amorphous structure due to the dielectric surfaces. The cladding layer **408** includes thickness ranging between 9 nm and 12 nm.

[0045] Turning to FIG. 5F, an anisotropic etching process, such as plasma etching, is applied to the cladding layer **408** to remove portions of the cladding layer **408** on the top surfaces of the semiconductor fins **218** and the top surface of the isolation features **230**. The anisotropic etching process includes suitable chemical, such as HBr, O₂, N₂, F or a combination thereof. Such formed silicon nitride layer **406**, the dielectric layer **404** and the cladding layer **408** are collectively referred to as the engineered cladding layer **220**.

[0046] Some embodiments are described with reference to FIGS. 6A-6G. The method may include operations described in FIG. 1C. The method is similar to the method associated with the structure in FIGS. 5A-5F except for the dielectric layer **404** is removed before forming the cladding layer **408**. The similar descriptions are not repeated herein for simplicity. Turning to FIG. 6A, a liner of a semiconductor material **402** is formed on sidewalls of the semiconductor fins **218** (such as **218A** and **218B**) and the trenches **219**. The liner **402** includes a semiconductor material, such as silicon. The liner **402** is formed by a suitable method, such as epitaxial growth.

[0047] Turning to FIG. 6B, isolation features **230**, such as STI features, are formed in the trenches **219**. In the disclosed embodiment, STI features are formed by deposition, CMP and etch.

[0048] Turning to FIG. 6C, a dielectric layer **404**, such as silicon oxide layer, is formed over the sidewalls of the semiconductor fins **218** in the trenches **219**. In the disclosed embodiment, the dielectric layer **404** is directly formed on the liner **402**.

[0049] Turning to FIG. 6D, a RTN process is applied to the workpiece **200**, thereby partially or completely converting the liner **402** of silicon into a silicon nitride layer **406**. The silicon nitride layer **406** functions as a diffusion blocking layer to prevent the diffusion of the subsequent formed cladding layer into the semiconductor layer stack **205**. The RTN process is implemented at an elevated temperature with a precursor having nitrogen-containing chemical, such as N.sub.2, ammonia (NH₃), other suitable gas or a combination thereof. The RTN process may include a carry gas, such as argon. The elevated temperature may range between 700° C. and 900° C. The silicon nitride layer may have a thickness ranging between 5 Angstrom and 20 Angstrom. In some embodiments, the partial pressure of the nitrogen-containing chemical ranges between 5 torr and 500 torr during the RTN process and the RTN process duration ranges between 30 sec and 60 sec.

[0050] Turning to FIG. 6E, the dielectric layer **404** is removed by etching so that the silicon nitride layer **406** is exposed. The process to remove the dielectric layer **404** includes performing a chemical oxide removal (COR) process; and performing a post heating treatment (PHT). The COR process includes applying NH_4F and HF at a first temperature **T1**; and the PHT includes applying an annealing process at a second temperature **T2** greater than the first temperature **T1**. In some embodiments, **T1** ranges between 10° C. and 100° C.; and **T2** ranges between 120° C. and 250° C. During the COR process, NH_4F and HF are absorbed on the surface of the dielectric layer **404**, causing a reaction that generates a byproduct, such as $(\text{NH}_4)_2\text{SiF}_6$. During the PHT process, the byproduct generated at COR is evaporated.

[0051] Turning to FIG. 6F, a cladding layer **408** of a semiconductor material, such as silicon germanium, is formed on the silicon nitride layer **406**. The cladding layer **408** and the liner **402** are different in composition. The cladding layer **408** may include a semiconductor material similar to that of the first semiconductor layers **210** in composition. In the depicted embodiment, the cladding layer **408** includes silicon germanium. The cladding layer **408** provides paths to etch the first semiconductor layers **210** and is removed with the first semiconductor layers **210** during a channel-release process at later stage. In the present embodiment, the cladding layer **408** is a semiconductor layer in an amorphous structure due to the dielectric surfaces.

[0052] Turning to FIG. 6G, an anisotropic etching process, such as plasma etching, may be applied to the cladding layer **408** to remove portions of the cladding layer **408** on the top surfaces of the semiconductor fins **218** and the top surface of the isolation features **230**. The anisotropic etching process includes suitable chemical, such as HBr, O_2 , N_2 , F or a combination thereof. Such formed silicon nitride layer **406** and the cladding layer **408** are collectively referred to as the engineered cladding layer **220**.

[0053] In an alternative embodiment illustrated in FIG. 6H, semiconductor fins **218** extends down into the substrate **202** and the top surface of the STI features **230** is below the semiconductor layer stack **205**. Thus, the cladding layer **408** includes different portions with different structures. Particularly, the sidewalls of the semiconductor fins **218** include top portions with dielectric surfaces (the surfaces of the silicon nitride layer **406**) and of the bottom portions with semiconductor surfaces (the surfaces of the liner **402**). The liner **402** is epitaxially grown on the surfaces of the semiconductor layer stack **205** and the substrate **202**, therefore has crystalline structure while the silicon nitride layer **406** has an amorphous structure. Accordingly, the portions **408A** of the cladding layer **408** grown on the silicon nitride layer **406** have an amorphous structure and the portions **408B** of the cladding layer **408** grown on the liner **402** have a crystalline structure.

[0054] Some embodiments are described with reference to FIGS. 7A-7G. The method may include operations described in FIG. 1D. Similar descriptions are not repeated for simplicity. Turning to FIG. 7A, a liner of a semiconductor material **402** is formed on sidewalls of the semiconductor fins **218** (such as **218A** and **218B**) and the trenches **219**. The liner **402** includes a semiconductor material, such as silicon. The liner **402** is formed by a suitable method, such as epitaxial growth.

[0055] Turning to FIG. 7B, isolation features **230**, such as STI features, are formed in the trenches **219**. In the disclosed embodiment, STI features are formed by deposition, CMP and etch.

[0056] Turning to FIG. 7C, a dielectric layer **404**, such as silicon oxide layer, is formed over the sidewalls of the semiconductor fins **218** in the trenches **219**. In the disclosed embodiment, the dielectric layer **404** is formed on the liner **402**.

[0057] Turning to FIG. 7D, an in-situ carbon (ISC) process is applied to the workpiece **200**, thereby forming a carbon-rich layer **410** on the dielectric layer **404**. The carbon-rich layer **410** is also referred to as carbon-rich oxide layer and is a dielectric layer that includes carbon, oxygen and nitrogen with high carbon concentration. The carbon-rich layer **410** and the dielectric layer **404** function as a collective diffusion blocking layer to prevent the diffusion of the subsequent formed cladding layer into the semiconductor layer stack **205**. The ISC process is implemented at an elevated temperature with a precursor having carbon-containing chemical, nitrogen-containing

chemical, carbon and nitrogen-containing chemical or a combination thereof. In some embodiments, the precursor includes $\text{H.sub.2Si[N(C.sub.2H.sub.5).sub.2].sub.2}$ with a chemical structure illustrated in FIG. 7G. During the ISC process, carbon and nitrogen in the precursor are not chemically broken, thereby forming a carbon- and nitrogen-containing dielectric layer (or carbon-rich nitride layer) **410**. In some embodiments, the precursor includes carbon-containing chemical (such as carbon dioxide (CO.sub.2), methane (CH.sub.4), other suitable chemical or a combination thereof) and a nitrogen-containing chemical (such as N.sub.2 , ammonia (NH_3), other suitable gas or a combination thereof). The ISC process may include a carry gas, such as argon and may be implemented at an elevated temperature. The carbon-rich layer **410** may have a thickness ranging between 10 Angstrom and 30 Angstrom.

[0058] In some embodiments, the dielectric layer **404** and the carbon-rich layer **410** may be sequentially formed in a same tool, such as in a same chamber of a CVD tool. In furtherance of the embodiments, the dielectric layer **404** is a undoped silica glass (USG) and the carbon-rich layer **410** is a carbon-nitrogen-containing dielectric layer. The formation of these two dielectric layers includes two steps. In the first step, the applied gases include

$\text{H.sub.2Si[N(C.sub.2H.sub.5).sub.2].sub.2}$, oxygen (O_2), and carry gas argon to form a USG layer. In the second step, the gas oxygen is turned off, and the applied gases include

$\text{H.sub.2Si[N(C.sub.2H.sub.5).sub.2].sub.2}$ and carry gas argon to form a carbon-nitrogen-containing dielectric layer. During the process, a radiofrequency (RF) power is applied to generate plasma, such as a RF power ranging between 500 W and 1000 W according to some examples. In some embodiments, the method to form the dielectric layer **404** and the carbon-rich layer **410** is a cyclic process that includes multiple cycles, each cycle including the first step and second step, thus forming a collective diffusion blocking layer. Especially, the number “n” of cycles, the first duration t_1 of the first step and the second duration t_2 of the second step in each cycle are tunable parameters optimized to form the collective blocking layer with enhanced protection effect.

[0059] Turning to FIG. 7E, a cladding layer **408** of a semiconductor material, such as silicon germanium, is formed on the carbon-rich layer **410** and the dielectric layer **404**. The cladding layer **408** and the liner **402** are different in composition. The cladding layer **408** may include a semiconductor material similar to that of the first semiconductor layers **210** in composition. In the depicted embodiment, the cladding layer **408** includes silicon germanium. The cladding layer **408** provides paths to etch the first semiconductor layers **210** and is removed with the first semiconductor layers **210** during a channel-release process at later stage. In the present embodiment, the cladding layer **408** is a semiconductor layer in an amorphous structure due to the dielectric surfaces. The cladding layer **408** includes thickness ranging between 50 Angstrom and 100 Angstrom.

[0060] Turning to FIG. 7F, an anisotropic etching process, such as plasma etching, is applied to the cladding layer **408** to remove portions of the cladding layer **408** on the top surfaces of the semiconductor fins **218** and the top surface of the isolation features **230**. The anisotropic etching process includes suitable chemical, such as HBr, O_2 , N_2 , F or a combination thereof. Such formed carbon-rich layer **410**, the dielectric layer **404** and the cladding layer **408** are collectively referred to as the engineered cladding layer **220**.

[0061] Some embodiments are described with reference to FIGS. 8A-8D. The method may include operations described in FIG. 1E. Similar descriptions are not repeated for simplicity. Turning to FIG. 8A, a liner of a semiconductor material **402** is formed on sidewalls of the semiconductor fins **218** (such as **218A** and **218B**) and the trenches **219**. The liner **402** includes a semiconductor material, such as silicon. The liner **402** is formed by a suitable method, such as epitaxial growth.

[0062] Turning to FIG. 8B, isolation features **230**, such as STI features, are formed in the trenches **219**. In the disclosed embodiment, STI features are formed by deposition, CMP and etch.

[0063] Turning to FIG. 8C, an etching process is applied to remove exposed portions of the liner **402** and oxide present on the surfaces of the semiconductor fins **218**, which provides improved

surface for selective epitaxial growth in the subsequent operation. The etching process may include multiple etching steps with various etchants to selectively remove Si liner and various oxide materials present on the surfaces of the semiconductor fins **218**. In some embodiments, the etching process includes an etching step using KOH solution for silicon etch and an etching step using buffered hydrofluoric acid (BHF) for oxide removal. In other examples, cryogenic deep reactive ion etching (DRIE) with SF₆ and O₂ is used to etch Si liner **402**.

[0064] Turning to FIG. **8D**, a cladding layer **408** of a semiconductor material, such as silicon germanium, is selectively formed on the semiconductor surface of semiconductor fins **218**. The cladding layer **408** may include a semiconductor material similar to that of the first semiconductor layers **210** in composition. In the depicted embodiment, the cladding layer **408** includes silicon germanium. The cladding layer **408** is formed by selective epitaxial growth, which is achieved by improved deposition surfaces and the tuned deposition parameters of the selective epitaxial growth. In the present embodiment, the cladding layer **408** is a semiconductor layer in a crystalline structure. The cladding layer **408** includes thickness ranging between 8 nm and 10 nm. Since the cladding layer **408** is selectively formed on the sidewalls of the semiconductor fins **218**, the anisotropic etching process is avoided. Furthermore, due to the selective growth, the footing issues of the cladding layer **408** is eliminated or reduced. In some embodiments, the SEG process is implemented with a precursor having Si-containing and Ge-containing chemicals, such as SiH₄, Si₂H₆, GeH₄, other suitable gas, or a combination thereof. In furtherance of the embodiments, the SEG process is implemented at an elevated temperature and a controlled chamber pressure. For example, the SEG process includes an elevated temperature ranging between 450° C. and 550° C., and a pressure of 1 torr or less.

[0065] As described above in accordance with various embodiments, the engineered cladding layer **220** may have different compositions and different structures, such as one illustrated in FIG. **5F**, one illustrated in FIG. **6G**, one illustrated in FIG. **6H**, or one illustrated in FIG. **7F**. In other embodiment, the engineered cladding layer **220** only includes the semiconductor cladding layer **408** of silicon germanium, as illustrated in FIG. **8D**. In the following descriptions, the engineered cladding layer **220** is only illustrated in figures and it is understood that the engineered cladding layer **220** is a composite layer and may include multiple films of different materials.

[0066] In some embodiments, dielectric fins **222** may be formed among the fins **218**. The dielectric fins **222** are dielectric features of one or more dielectric material. Only one dielectric fin **222** is illustrated in FIG. **9B**. More dielectric fins **222** may be present, such as one on left side of the fin **218A** and another one on right side of the fin **218B**. Dielectric fin **222** may be formed by any suitable method that including deposition. In some embodiments, dielectric fin **222** includes a dielectric stack **222A** and a self-aligned cap **222B** disposed on the dielectric stack **222A** and aligned with the dielectric stack **222A**, as illustrated in FIG. **9E**. In furtherance of the embodiment, the dielectric fin **222** is formed by a procedure that includes a deposition of one or more dielectric material (such as silicon oxide or other suitable dielectric material) to fill in the gap between the fins **218**; performing a chemical mechanical polishing (CMP) process; selectively etching to recess the deposited dielectric material; depositing another dielectric material (such as high-k dielectric material, e.g., hafnium nitride, metal oxide, metal nitride, metal oxynitride or a combination thereof) and performing another CMP process to form the dielectric stack **222A** and the self-aligned cap **222B**. In some embodiments, the dielectric fin **222** includes a conformal dielectric layer **222C** and a bulk dielectric layer **222D** disposed on the conformal dielectric layer **222C**, as illustrated in FIG. **9F**. In furtherance of the embodiment, the dielectric fin **222** is formed by a procedure that includes a conformal deposition of one or more dielectric material in the gap between the fins **218** and depositing another dielectric material on the conformal dielectric layer **222C** to fill in the gap between the fins **218**; and performing a CMP process. In some embodiments, the hard mask used to pattern semiconductor stacks **205** may be removed at this stage. Thus, the dielectric fin **222** is extended above the fins **218**.

[0067] Turning to FIGS. 10A-10D, gate structures 240 are formed over portions of fins 218A, 218B, over dielectric fin 222, and over isolation features 230. Gate structures 240 extend lengthwise in a direction that is different than (e.g., orthogonal to) the lengthwise direction of fins 218A, 218B. For example, gate structures 240 extend substantially parallel to one another along the x-direction, having a length defined in the y-direction, a width defined in the x-direction, and a height defined in the z-direction. Gate structures 240 are disposed on portions of fins 218A, 218B and define source/drain regions 242 and channel regions 244 of fins 218A, 218B. In the X-Z plane, gate structures 240 wrap top surfaces and sidewall surfaces of fins 218A, 218B. In the Y-Z plane, gate structures 240 are disposed over top surfaces of respective channel regions 244 of fins 218A, 218B, such that gate structures 240 interpose respective source/drain regions 242. Each gate structure 240 includes a gate region 240-1 that corresponds with a portion of the respective gate structure 240 that will be configured for an n-type GAA transistor (and thus corresponds with a portion spanning an n-type GAA transistor region) and a gate region 240-2 that corresponds with a portion of the respective gate structure 240 that will be configured for a p-type GAA transistor (and thus corresponds with a portion spanning a p-type GAA transistor region). Gate structures 240 may be configured differently in gate region 240-1 and gate region 240-2, depending on the transistors to be formed on these regions, such as p-type transistors or n-type transistors. For example, each of gate structures 240 spans gate region 240-1 and gate region 240-2 and may be configured differently in gate region 240-1 and gate region 240-2 to optimize performance of the n-type GAA transistors (having n-gate electrodes in gate regions 240-1) and the p-type GAA transistors (having p-gate electrodes in gate regions 240-2). Accordingly, gate regions 240-1 will be referred to as n-type gate regions 240-1 and gate regions 240-2 will be referred to as p-type gate regions 240-2 hereinafter.

[0068] In FIGS. 10A-10D, each gate structure 240 includes a dummy gate stack 245. In the depicted embodiment, a width of dummy gate stacks 245 defines a gate length ($L_{sub.g}$) of gate structures 240 (here, in the y-direction), where the gate length defines a distance (or length) that current (e.g., carriers, such as electrons or holes) travels between source/drain regions 242 when the n-type GAA transistor and/or the p-type GAA transistor are switched (turned) on. In some embodiments, the gate length is about 5 nm to about 250 nm. Gate length can be tuned to achieve desired operation speeds of the GAA transistors and/or desired packing density of the GAA transistors. For example, when a GAA transistor is switched on, current flows between source/drain regions of the GAA transistor. Increasing the gate length increases a distance required for current to travel between the source/drain regions, increasing a time it takes for the GAA transistor to switch fully on. Conversely, decreasing the gate length decreases the distance required for current to travel between the source/drain regions, decreasing a time it takes for the GAA transistor to switch fully on. Smaller gate lengths provide GAA transistors that switch on/off more quickly, facilitating faster, high speed operations. Smaller gate lengths also facilitate tighter packing density (i.e., more GAA transistors can be fabricated in a given area of an IC chip), increasing a number of functions and applications that can be fabricated on the IC chip. In the depicted embodiment, the gate length of one or more of gate structures 240 is configured to provide GAA transistors having short-length channels. For example, the gate length of GAA transistors is about 5 nm to about 20 nm. In some embodiments, multigate device 200 can include GAA transistors having different gate lengths.

[0069] Dummy gate stacks 245 include a dummy gate electrode, and in some embodiments, a dummy gate dielectric. The dummy gate electrode includes a suitable dummy gate material, such as polysilicon layer. In embodiments where dummy gate stacks 245 include a dummy gate dielectric disposed between the dummy gate electrode and fins 218A, 218B, the dummy gate dielectric includes a dielectric material, such as silicon oxide, a high-k dielectric material, other suitable dielectric material, or combinations thereof. Examples of high-k dielectric material include HfO_{2} , $HfSiO$, $HfSiON$, $HfTaO$, $HfTiO$, $HfZrO$, zirconium oxide, aluminum oxide, hafnium dioxide-alumina (HfO_{2} — $Al_{2}O_{3}$) alloy, other suitable high-k dielectric materials, or

combinations thereof. In some embodiments, the dummy gate dielectric includes an interfacial layer (including, for example, silicon oxide) disposed over fins **218A**, **218B** and a high-k dielectric layer disposed over the interfacial layer. Dummy gate stacks **245** can include numerous other layers, for example, capping layers, interface layers, diffusion layers, barrier layers, hard mask layers, or combinations thereof. For example, dummy gate stacks **245** can further include a hard mask layer disposed over the dummy gate electrode.

[0070] Dummy gate stacks **245** are formed by deposition processes, lithography processes, etching processes, other suitable processes, or combinations thereof. For example, a deposition process is performed to form a dummy gate electrode layer over fins **218A**, **218B** and isolation features **230**. In some embodiments, a deposition process is performed to form a dummy gate dielectric layer over fins **218A**, **218B** and isolation features **230** before forming the dummy gate electrode layer. In such embodiments, the dummy gate electrode layer is deposited over the dummy gate dielectric layer. In some embodiment, a hard mask layer is deposited over the dummy gate electrode layer. The deposition process includes CVD, physical vapor deposition (PVD), atomic layer deposition (ALD), high density plasma CVD (HDPCVD), metal organic CVD (MOCVD), remote plasma CVD (RPCVD), plasma enhanced CVD (PECVD), low-pressure CVD (LPCVD), atomic layer CVD (ALCVD), atmospheric pressure CVD (APCVD), plating, other suitable methods, or combinations thereof. A lithography patterning and etching process is then performed to pattern the dummy gate electrode layer (and, in some embodiments, the dummy gate dielectric layer and the hard mask layer) to form dummy gate stacks **245**, such that dummy gate stacks **245** (including the dummy gate electrode layer, the dummy gate dielectric layer, the hard mask layer, and/or other suitable layers) is configured as depicted in FIGS. **10A-10D**. The lithography patterning processes include resist coating (for example, spin-on coating), soft baking, mask aligning, exposure, post-exposure baking, developing the resist, rinsing, drying (for example, hard baking), other suitable lithography processes, or combinations thereof. The etching processes include dry etching processes, wet etching processes, other etching methods, or combinations thereof.

[0071] In some embodiments, a lightly doped source/drain (LDD) implantation process may be applied to the semiconductor layer stack **205** to form LDD features (not shown) aligned with edges of the dummy gate stacks **245**. LDD features are separately formed for n-type GAA transistors and p-type GAA transistors. For example, LDD features for n-type GAA transistors includes n-type dopant, such as phosphorous while LDD features for p-type GAA transistors includes p-type dopant, such as boron.

[0072] Each gate structure **240** further includes gate spacers **247** disposed adjacent to (i.e., along sidewalls of) respective dummy gate stacks **245**. Gate spacers **247** are formed by any suitable process and include a dielectric material. The dielectric material can include silicon, oxygen, carbon, nitrogen, other suitable material, or combinations thereof (e.g., silicon oxide, silicon nitride, silicon oxynitride (SiON), silicon carbide, silicon carbon nitride (SiCN), silicon oxycarbide (SiOC), silicon oxycarbon nitride (SiOCN)). For example, a dielectric layer including silicon and nitrogen, such as a silicon nitride layer, can be deposited over dummy gate stacks **245** and subsequently etched (e.g., anisotropically etched) to form gate spacers **247**. In some embodiments, gate spacers **247** include a multi-layer structure, such as a first dielectric layer that includes silicon nitride and a second dielectric layer that includes silicon oxide. In some embodiments, more than one set of spacers, such as seal spacers, offset spacers, sacrificial spacers, dummy spacers, and/or main spacers, are formed adjacent to dummy gate stacks **245**. In such implementations, the various sets of spacers can include materials having different etch rates. For example, a first dielectric layer including silicon and oxygen (e.g., silicon oxide) can be deposited and etched to form a first spacer set adjacent to dummy gate stacks **245**, and a second dielectric layer including silicon and nitrogen (e.g., silicon nitride) can be deposited and etched to form a second spacer set adjacent to the first spacer set. Note that the workpiece **200** illustrated in FIGS. **10A-10D** and subsequent figures includes more dielectric fins **222**. It is not intended to be limiting, and more or less dielectric fins

222 may be present according to different embodiments.

[0073] Turning to FIGS. 11A-11D, exposed portions of fins **218A**, **218B** (i.e., source/drain regions **242** of fins **218A**, **218B** that are not covered by gate structures **240**) are at least partially removed to form source/drain trenches (recesses) **250**. In the depicted embodiment, an etching process completely removes semiconductor layer stack **205** in source/drain regions **242** of fins **218A**, **218B**, thereby exposing the substrate portion of fins **218A**, **218B** in source/drain regions **242** (e.g., p-well **204A** and n-well **204B**). Source/drain trenches **250** thus have sidewalls defined by remaining portions of semiconductor layer stack **205**, which are disposed in channel regions **244** under gate structures **240**, and bottoms defined by substrate **202**, such as top surfaces of p-well **204A** and n-well **204B** in source/drain regions **242**. In some embodiments, the etching process removes some, but not all, of semiconductor layer stack **205**, such that source/drain trenches **250** have bottoms defined by semiconductor layer **210** or semiconductor layer **215** in source/drain regions **242**. In some embodiments, the etching process further removes some, but not all, of the substrate portion of fins **218A**, **218B**, such that source/drain recesses **250** extend below a topmost surface of substrate **202**. The etching process can include a dry etching process, a wet etching process, other suitable etching process, or combinations thereof. In some embodiments, the etching process is a multi-step etch process. For example, the etching process may include alternative etchants to separately and alternately remove semiconductor layers **210** and semiconductor layers **215**. In some embodiments, parameters of the etching process are configured to selectively etch semiconductor layer stack with minimal (to no) etching of gate structures **240** (i.e., dummy gate stacks **245** and gate spacers **247**) and/or isolation features **230**. In some embodiments, a lithography process, such as those described herein, is performed to form a patterned mask layer that covers gate structures **240** and/or isolation features **230**, and the etching process uses the patterned mask layer as an etch mask. In some embodiments, the patterned mask layer used to form dummy gate stacks **245** remains and is used as an etch mask to recess the source/drain regions **242**.

[0074] Turning to FIGS. 12A-12D, inner spacers **255** are formed in channel regions **244** along sidewalls of semiconductor layers **210** by any suitable process. Particularly, the inner spacers **255** are formed to be vertically aligned with gate spacers **247** (and the LDD features if present) to provide isolation and separation between the gate structure and the source/drain features. For example, a first etching process is performed that selectively etches semiconductor layers **210** exposed by source/drain trenches **250** with minimal (to no) etching of semiconductor layers **215**, such that gaps are formed between semiconductor layers **215** and between semiconductor layers **215** and substrate **202** under gate spacers **247**. Portions (edges) of semiconductor layers **215** are thus suspended in the channel regions **244** under gate spacers **247**. In some embodiments, the gaps extend partially under dummy gate stacks **245**. The first etching process is configured to laterally etch (e.g., along the y-direction) semiconductor layers **210**, thereby reducing a length of semiconductor layers **210** along the y-direction. The first etching process is a dry etching process, a wet etching process, other suitable etching process, or combinations thereof. A deposition process then forms a spacer layer over gate structures **240** and over features defining source/drain trenches **250** (e.g., semiconductor layers **215**, semiconductor layers **210**, and substrate **202**), such as CVD, PVD, ALD, HDPCVD, MOCVD, RPCVD, PECVD, LPCVD, ALCVD, APCVD, plating, other suitable methods, or combinations thereof. The spacer layer partially (and, in some embodiments, completely) fills the source/drain trenches **250**. The deposition process is configured to ensure that the spacer layer fills the gaps between semiconductor layers **215** and between semiconductor layers **215** and substrate **202** under gate spacers **247**. A second etching process is then performed that selectively etches the spacer layer to form inner spacers **255** as depicted in FIGS. 12A-12D with minimal (to no) etching of semiconductor layers **215**, dummy gate stacks **245**, and gate spacers **247**. In some embodiments, the spacer layer is removed from sidewalls of gate spacers **247**, sidewalls of semiconductor layers **215**, dummy gate stacks **245**, and substrate **202**. The spacer layer (and thus inner spacers **255**) includes a material that is different than a material of semiconductor

layers **215** and a material of gate spacers **247** to achieve desired etching selectivity during the second etching process. In some embodiments, the spacer layer includes a dielectric material that includes silicon, oxygen, carbon, nitrogen, other suitable material, or combinations thereof (for example, silicon oxide, silicon nitride, silicon oxynitride, silicon carbide, or silicon oxycarbonitride). In some embodiments, the spacer layer includes a low-k dielectric material, such as those described herein. In some embodiments, dopants (for example, p-type dopants, n-type dopants, or combinations thereof) are introduced into the dielectric material, such that spacer layer includes a doped dielectric material.

[0075] Turning to FIGS. **13A-13D**, epitaxial source/drain features are formed in source/drain recesses **250**. For example, a semiconductor material is epitaxially grown from portions of substrate **202** and semiconductor layers **215** exposed by source/drain recesses **250**, forming epitaxial source/drain features **260A** in source/drain regions **242** that correspond with n-type GAA transistor regions and epitaxial source/drain features **260B** in source/drain regions **242** that correspond with p-type GAA transistor regions. An epitaxy process can use CVD deposition techniques (for example, VPE and/or UHV-CVD), molecular beam epitaxy, other suitable epitaxial growth processes, or combinations thereof. The epitaxy process can use gaseous and/or liquid precursors, which interact with the composition of substrate **202** and/or semiconductor layer stack **205** (in particular, semiconductor layers **215**). Epitaxial source/drain features **260A**, **260B** are doped with n-type dopants and/or p-type dopants. In some embodiments, for the n-type GAA transistors, epitaxial source/drain features **260A** include silicon. Epitaxial source/drain features **260A** can be doped with carbon, phosphorous, arsenic, other n-type dopant, or combinations thereof (for example, forming Si:C epitaxial source/drain features, Si:P epitaxial source/drain features, or Si:C:P epitaxial source/drain features). In some embodiments, for the p-type GAA transistors, epitaxial source/drain features **260B** include silicon germanium or germanium. P-type epitaxial source/drain features **260B** can be doped with boron, other p-type dopant, or combinations thereof (for example, forming Si:Ge:B epitaxial source/drain features). In some embodiments, epitaxial source/drain features **260A** and/or epitaxial source/drain features **260B** include more than one epitaxial semiconductor layer, where the epitaxial semiconductor layers can include the same or different materials and/or dopant concentrations. In some embodiments, epitaxial source/drain features **260A**, **260B** include materials and/or dopants that achieve desired tensile stress and/or compressive stress in respective channel regions **244**. In some embodiments, epitaxial source/drain features **260A**, **260B** are doped during deposition by adding impurities to a source material of the epitaxy process (i.e., in-situ). In some embodiments, epitaxial source/drain features **260A**, **260B** are doped by an ion implantation process subsequent to a deposition process. In some embodiments, annealing processes (e.g., rapid thermal annealing (RTA) and/or laser annealing) are performed to activate dopants in epitaxial source/drain features **260A**, **260B** and/or other source/drain regions (for example, heavily doped source/drain regions and/or lightly doped source/drain (LDD) regions). In some embodiments, epitaxial source/drain features **260A**, **260B** are formed in separate processing sequences that include, for example, masking p-type GAA transistor regions when forming epitaxial source/drain features **260A** in n-type GAA transistor regions and masking n-type GAA transistor regions when forming epitaxial source/drain features **260B** in p-type GAA transistor regions.

[0076] In some embodiments, the source/drain features **260A** and **260B** may be engineered to have desired shapes and sizes by tuning the etching process that forms the source/drain recesses **250** and the epitaxial growth to form the source/drain features **260A/260B**.

[0077] In some embodiments, the deposition (epitaxial growth) chemical in the precursor may include silane (SiH₄) or dichlorosilane (SiH₂Cl₂) for growing silicon, GeH₄ for growing germanium, or both for growing silicon germanium. The precursor also includes chemical for dopant, such as phosphorous-containing chemical for n-type dopant or boron-containing chemical for p-type dopant. In the present embodiment, the precursor for n-type source/drain

features **260A** includes SiH.sub.4 and a phosphorous-containing chemical to form the n-type source/drain features **260A** of silicon doped with phosphorous. The precursor for epitaxial growth may additionally include etching chemical to control the epitaxial growth and the profile of the source/drain features. In some embodiments, the etching chemical includes HCl. In some embodiments, the etching chemical includes chlorine-containing chemical, such as HCl or Cl.sub.2, or fluorine-containing chemical, such as SF.sub.6, or alternatively both chlorine-containing chemical and fluorine-containing chemical.

[0078] In some embodiments, the epitaxial growth is designed with a lower deposition/etching (D/E) ratio to achieve a desired source/drain profile, the etching gas in the precursor uses HCl with a flow rate greater than 30000 sccm, or a flow rate ranging between 40000 sccm~30000 sccm. Thus, the epitaxial growth is a bottom-up deposition, thereby forming the source/drain features with a geometry having substantial vertical sidewalls. In some embodiments, the epitaxial growth is designed with a higher ratio D/E to achieve a different profile, the etching gas in the precursor uses HCl with a flow rate less than 10000 sccm, or ranging between 0 sccm~10000 sccm, the epitaxial growth forms the source/drain features with a profile having uneven sidewalls.

[0079] Turning to FIGS. **14A-14D**, an inter-level dielectric (ILD) layer **270** is formed over isolation features **230**, epitaxial source/drain features **260A**, **260B**, and gate spacers **247**, for example, by a deposition process (such as CVD, PVD, ALD, HDPCVD, MOCVD, RPCVD, PECVD, LPCVD, ALCVD, APCVD, plating, other suitable methods, or combinations thereof). ILD layer **270** is disposed between adjacent gate structures **240**. In some embodiments, ILD layer **270** is formed by a flowable CVD (FCVD) process that includes, for example, depositing a flowable material (such as a liquid compound) over multigate device **200** and converting the flowable material to a solid material by a suitable technique, such as thermal annealing and/or ultraviolet radiation treating. ILD layer **270** includes a dielectric material including, for example, silicon oxide, silicon nitride, silicon oxynitride, TEOS formed oxide, PSG, BPSG, low-k dielectric material, other suitable dielectric material, or combinations thereof. Exemplary low-k dielectric materials include FSG, carbon doped silicon oxide, Black Diamond® (Applied Materials of Santa Clara, California), Xerogel, Aerogel, amorphous fluorinated carbon, Parylene, BCB, SiLK (Dow Chemical, Midland, Michigan), polyimide, other low-k dielectric material, or combinations thereof. In the depicted embodiment, ILD layer **270** is a dielectric layer that includes a low-k dielectric material (generally referred to as a low-k dielectric layer). ILD layer **270** can include a multilayer structure having multiple dielectric materials. In some embodiments, a contact etch-stop layer (CESL) is disposed between ILD layer **270** and isolation features **230**, epitaxial source/drain features **260A**, **260B**, and gate spacers **247**. The CESL includes a material different than ILD layer **270**, such as a dielectric material that is different than the dielectric material of ILD layer **270**. For example, where ILD layer **270** includes a low-k dielectric material, the CESL includes silicon and nitrogen, such as silicon nitride or silicon oxynitride. Subsequent to the deposition of ILD layer **270** and/or the CESL, a CMP process and/or other planarization process can be performed until reaching (exposing) a top portion (or top surface) of dummy gate stacks **245**. In some embodiments, the planarization process removes hard mask layers of dummy gate stacks **245** to expose underlying dummy gate electrodes of dummy gate stacks **245**, such as polysilicon gate electrode layers.

[0080] ILD layer **270** may be a portion of a multilayer interconnect (MLI) feature disposed over substrate **202**. The MLI feature electrically couples various devices (for example, p-type GAA transistors and/or n-type GAA transistors of multigate device **200**, transistors, resistors, capacitors, and/or inductors) and/or components (for example, gate structures and/or epitaxial source/drain features of p-type GAA transistors and/or n-type GAA transistors), such that the various devices and/or components can operate as specified by design requirements of multigate device **200**. The MLI feature includes a combination of dielectric layers and electrically conductive layers (e.g., metal layers) configured to form various interconnect structures. The conductive layers are configured to form vertical interconnect features, such as device-level contacts and/or vias, and/or

horizontal interconnect features, such as conductive lines. Vertical interconnect features typically connect horizontal interconnect features in different layers (or different planes) of the MLI feature. During operation, the interconnect features are configured to route signals between the devices and/or the components of multigate device **200** and/or distribute signals (for example, clock signals, voltage signals, and/or ground signals) to the devices and/or the components of multigate device **200**.

[0081] Turning to FIGS. **15A-15D**, dummy gate stacks **245** are removed from gate structures **240** by a suitable etching process, thereby resulting in gate trenches **275** and exposing semiconductor layer stacks **205** of fins **218A**, **218B** in n-type gate regions **240-1** and p-type gate regions **240-2**. The etching process is designed with etchant to selectively remove the dummy gate stacks **245**. In the depicted embodiment, an etching process completely removes dummy gate stacks **245** to expose semiconductor layers **215** and semiconductor layers **210** in channel regions **244**. The etching process is a dry etching process, a wet etching process, other suitable etching process, or combinations thereof. In some embodiments, the etching process is a multi-step etch process. For example, the etching process may include alternative etchants to separately remove various layers of dummy gate stacks **245**, such as the dummy gate electrode layers, the dummy gate dielectric layers, and/or the hard mask layers. In some embodiments, the etching process is configured to selectively etch dummy gate stacks **245** with minimal (to no) etching of other features of multigate device **200**, such as ILD layer **270**, gate spacers **247**, isolation features **230**, semiconductor layers **215**, and semiconductor layers **210**. In some embodiments, a lithography process, such as those described herein, is performed to form a patterned mask layer that covers ILD layer **270** and/or gate spacers **247**, and the etching process uses the patterned mask layer as an etch mask.

[0082] Turning to FIGS. **16A-16E**, cladding layer **208** and semiconductor layers **210** of semiconductor layer stack **205** (exposed by gate trenches **275**) are selectively removed from channel regions **244**, thereby forming suspended semiconductor layers **215** in channel regions **244**. In the depicted embodiment, an etching process selectively etches semiconductor layers **210** with minimal (to no) etching of semiconductor layers **215** and, in some embodiments, minimal (to no) etching of gate spacers **247** and/or inner spacers **255**. Various etching parameters can be tuned to achieve selective etching of semiconductor layers **210**, such as etchant composition, etching temperature, etching solution concentration, etching time, etching pressure, source power, RF bias voltage, RF bias power, etchant flow rate, other suitable etching parameters, or combinations thereof. For example, an etchant is selected for the etching process that etches the material of semiconductor layers **210** (in the depicted embodiment, silicon germanium) at a higher rate than the material of semiconductor layers **215** (in the depicted embodiment, silicon) (i.e., the etchant has a high etch selectivity with respect to the material of semiconductor layers **210**). The etching process is a dry etching process, a wet etching process, other suitable etching process, or combinations thereof. In some embodiments, a dry etching process (such as an RIE process) utilizes a fluorine-containing gas (for example, SF₆) to selectively etch semiconductor layers **210**. In some embodiments, a ratio of the fluorine-containing gas to an oxygen-containing gas (for example, O₂), an etching temperature, and/or an RF power may be tuned to selectively etch silicon germanium (or silicon). In some embodiments, a wet etching process utilizes an etching solution that includes ammonium hydroxide (NH₄OH) and water (H₂O) to selectively etch semiconductor layers **210**. In some embodiments, a chemical vapor phase etching process using hydrochloric acid (HCl) to selectively etch semiconductor layers **210**.

[0083] At least one suspended semiconductor layer **215** is thus exposed in n-type gate regions **240-1** and p-type gate regions **240-2** by gate trenches **275**. In the depicted embodiment, each n-type gate region **240-1** and each p-type gate region **240-2** includes four suspended semiconductor layers **215** vertically stacked that will provide four channels through which current will flow between respective epitaxial source/drain features (epitaxial source/drain features **260A** or epitaxial source/drain features **260B**) during operation of the GAA transistors. Suspended semiconductor

layers **215** are thus referred to as channel layers **215'** in n-type gate regions **240-1** are separated by gaps **277A**, and channel layers **215'** in p-type gate regions **240-2** are separated by gaps **277B**, collectively being referred to as gaps **277**. Channel layers **215'** in n-type gate regions **240-1** are also separated from substrate **202** by gaps **277A**, and channel layers **215'** in p-type gate regions **240-2** are also separated from substrate **202** by gaps **277B**. A spacing **s1** is defined between channel layers **215'** along the z-direction in n-type gate regions **240-1**, and a spacing **s2** is defined between channel layers **215'** along the z-direction in p-type gate regions **240-2**. Spacing **s1** and spacing **s2** correspond with a width of gaps **277A** and gaps **277B**, respectively. In the depicted embodiment, spacing **s1** is about equal to **s2**, though the present disclosure contemplates embodiments where spacing **s1** is different than spacing **s2**. In some embodiments, spacing **s1** and spacing **s2** are both about equal to thickness **t1** of semiconductor layers **210**. Further, channel layers **215'** in n-type gate regions **240-1** have a length **l1** along the x-direction and a width **w1** along the z-direction, and channel layers **215'** in p-type gate regions **240-2** have a length **l2** along the x-direction and a width **w2** along the z-direction. In the depicted embodiment, length **l1** is about equal to length **l2**, and width **w1** is about equal to width **w2**, though the present disclosure contemplates embodiments where length **l1** is different than length **l2** and/or width **w1** is different than width **w2**. In some embodiments, length **l1** and/or length **l2** is about 10 nm to about 50 nm. In some embodiments, width **w1** and/or width **w2** is about 4 nm to about 10 nm. In some embodiments, each channel layer **215'** has nanometer-sized dimensions and can be referred to as a "nanowire," which generally refers to a channel layer suspended in a manner that will allow a metal gate to physically contact at least two sides of the channel layer, and in GAA transistors, will allow the metal gate to physically contact at least four sides of the channel layer (i.e., surround the channel layer). In such embodiments, a vertical stack of suspended channel layers can be referred to as a nanostructure, and the process depicted in FIGS. **16A-16D** can be referred to as a channel nanowire release process. In some embodiments, after removing semiconductor layers **210**, an etching process is performed to modify a profile of channel layers **215'** to achieve desired dimensions and/or desired shapes (e.g., cylindrical-shaped (e.g., nanowire), rectangular-shaped (e.g., nanobar), sheet-shaped (e.g., nanosheet, etc., such as illustrated in FIG. **16B**). The present disclosure further contemplates embodiments where the channel layers **215'** (nanowires, as illustrated in FIG. **16E**) have sub-nanometer dimensions depending on design requirements of multigate device **200**. In furtherance of the embodiment where the channel layers **215'** are nanowires, the length along x-direction and width along z-direction are substantially equal. For example, **l1=w1** and **l2=w2** with relative difference being less than 10%.

[0084] Turning to FIGS. **17A-17E**, gate stacks **360A** (for n-type transistors) and **360B** (for p-type transistors) are formed over multigate device **200**. Gate stacks **360A** and **360B** are collectively referred to as gate stacks **360**. The gate stacks **360** are formed in the gate trenches **275** and are extended down to wrap around each of the vertically stacked channel layers **215'**.

[0085] The formation of the gate stacks includes deposition and planarization process, such as CMP. The gate stacks **360A** and **360B** may be collectively formed or alternatively, separately formed, depending on the type of GAA transistors, such as n-type GAA transistors or p-type GAA transistors. Accordingly, the gate stacks **360A** and **360B** may have the same compositions or alternatively different compositions, such as different work function metal layers (as described below). Each of the gate stacks **360A** and **360B** includes a gate dielectric layer and a gate electrode disposed on the gate dielectric layer. In some embodiments, the gate dielectric layer includes an interfacial layer **280** and a high-k dielectric layer disposed on the interfacial layer **280**. The gate electrode may include one or more conductive materials, such as a capping layer, a work function metal layer, a blocking layer, a metal fill layer, and/or other proper conductive material layers. In some embodiments, the gate electrode includes a work function layer (such as the work function layer **300** for the gate stack **360A** or the work function layer **310** for the gate stack **360B**) and a metal fill layer **350** disposed on the work function metal layer. The work function layers **300** and

310 may be same or different and may be an n-type work function layer or a p-type work function layer, depending on the types of the corresponding GAA transistors.

[0086] In the depicted embodiment, the gate dielectric layer includes an interfacial layer **280** and a high-k dielectric layer **282**, where interfacial layer **280** is disposed between the high-k dielectric layer **282** and channel layers **215'**. In furtherance of the depicted embodiment, interfacial layer **280** and high-k dielectric layer **282** partially fill gaps **277A** between channel layers **215'** and between channel layers **215'** and substrate **202** in the first gate region **240-1** and partially fill gaps **277B** between channel layers **215'** and between channel layers **215'** and substrate **202** in the second gate region **240-2**. In some embodiments, interfacial layer **280** and/or high-k dielectric layer **282** are also disposed on substrate **202**, isolation features **230**, and/or gate spacers **247**. Interfacial layer **280** includes a dielectric material, such as SiO₂, HfSiO, SiON, other silicon-comprising dielectric material, other suitable dielectric material, or combinations thereof. High-k dielectric layer **282** includes a high-k dielectric material, such as HfO₂, HfSiO, HfSiO₂, HfSiON, HfLaO, HfTaO, HfTiO, HfZrO, HfAlO_x, ZrO, ZrO₂, ZrSiO₂, AlO, AlSiO, Al₂O₃, TiO, TiO₂, LaO, LaSiO, Ta₂O₃, Ta₂O₅, Y₂O₃, SrTiO₃, BaZrO, BaTiO₃ (BTO), (Ba,Sr) TiO₃ (BST), Si₃N₄, hafnium dioxide-alumina (HfO₂—Al₂O₃) alloy, other suitable high-k dielectric material, or combinations thereof. High-k dielectric material generally refers to dielectric materials having a high dielectric constant, for example, greater than that of silicon oxide ($k \approx 3.9$). Interfacial layer **280** is formed by any of the processes described herein, such as thermal oxidation, chemical oxidation, ALD, CVD, other suitable process, or combinations thereof. In some embodiments, interfacial layer **280** has a thickness of about 0.5 nm to about 3 nm. High-k dielectric layer **282** is formed by any of the processes described herein, such as ALD, CVD, PVD, oxidation-based deposition process, other suitable process, or combinations thereof. In some embodiments, high-k dielectric layer **282** has a thickness of about 1 nm to about 2 nm.

[0087] The work function layer (**300** or **310**) is formed over multigate device **200**, particularly over high-k dielectric layer **282**. For example, an ALD process conformally deposits the work function layer on high-k dielectric layer **282**, such that the work function layer has a substantially uniform thickness and partially fills gate trenches **275**. The work function layer can be formed using another suitable deposition process, such as CVD, PVD, HDPCVD, MOCVD, RPCVD, PECVD, LPCVD, ALCVD, APCVD, spin coating, plating, other deposition process, or combinations thereof. For example, the work function layer is disposed along sidewalls, tops, and bottoms of channel layers **215'**. A thickness of the work function layer is configured to at least partially fill gaps (**277A** or **277B**) between channel layers **215'** and between channel layers **215'** and substrate **202** (and, in some embodiments, without filling gate trenches **275** along the gate length direction (here, along the y-direction)). In some embodiments, the work function layer has a thickness of about 1 nm to about 10 nm. In some embodiments, p-type work function layer includes any suitable p-type work function material, such as TiN, TaN, TaSN, Ru, Mo, Al, WN, WCN, ZrSi₂, MoSi₂, TaSi₂, NiSi₂, other p-type work function material, or combinations thereof. In the depicted embodiment, p-type work function layer includes titanium and nitrogen, such as TiN. In some embodiments, the n-type work function layer includes any suitable n-type work function material, such as Ti, Al, Ag, Mn, Zr, TiAl, TiAlC, TiAlSiC, TaC, TaCN, TaSiN, TaAl, TaAlC, TaSiAlC, TiAlN, other n-type work function material, or combinations thereof. In the depicted embodiment, n-type work function layer includes aluminum.

[0088] A metal fill (or bulk) layer **350** is formed over multigate device **200**, particularly over the work function layer **300** in the first gate regions **240-1** and over the work function layer **310** in the second gate regions **240-2**. For example, a CVD process or a PVD process deposits metal fill layer **350** such that metal fill layer **350** fills any remaining portion of gate trenches **275**, including any remaining portions of gaps (**277A** or **277B**) in the gate regions **240-1** and **240-2**. Metal fill layer **350** includes a suitable conductive material, such as Al, W, and/or Cu. Metal fill layer **350** may

additionally or collectively include other metals, metal oxides, metal nitrides, other suitable materials, or combinations thereof. Alternatively, metal fill layer **350** is formed using another suitable deposition process, such as ALD, CVD, PVD, HDPCVD, MOCVD, RPCVD, PECVD, LPCVD, ALCVD, APCVD, spin coating, plating, other deposition process, or combinations thereof. In the depicted embodiment, the metal fill layer **350** is formed by PVD to form a seed layer and followed by plating to completely form the metal fill layer.

[0089] A planarization process is performed to remove excess gate materials from multigate device **200**. For example, a CMP process is performed until a top surface of ILD layer **270** is reached (exposed), such that a top surface of gate structures **240** are substantially planar with a top surface of ILD layer **270** after the CMP process. Accordingly, multigate device **200** includes first GAA transistors having a gate stack **360A** wrapping respective channel layers **215'**, such that gate stack **360A** is disposed between respective epitaxial source/drain features **260A**, and second GAA transistor having metal gate stacks **360B** wrapping respective channel layers **215'**, such that metal gate stacks **360B** are disposed between respective epitaxial source/drain features **260B**.

[0090] Other embodiments may present. As described above, the channel layers **215'** of multigate device **200** may have different structure, such as nanosheet illustrated in FIG. **17B** or nanowire, such as illustrated in FIG. **17E**. In FIG. **17E**, the vertical dimension w and horizontal dimension between source/drain features of the channel layer **215'** are substantially equal, such $l1=w1$ and $l2=w2$.

[0091] The engineered cladding layer **220** is a composite layer and includes different compositions, such as a semiconductor film **408** (e.g., SiGe) and one or more dielectric films (such as dielectric layer **404**, silicon nitride layer **406**, carbon-rich oxide layer **410**, or a combination thereof). As described above according to some embodiments, the engineered cladding layer **220** is removed during the channel-release process, which may be achieved by multiple etching steps with etchants to remove respective materials. In some alternative embodiments, the engineered cladding layer **220** is partially removed during the channel-release process. For example, the etching process selectively removes the first semiconductor layers **210** and the cladding layer **408** as both have silicon germanium. The remained portion of the engineered cladding layer **220** after the channel-release process is illustrated in FIGS. **18A-18E** and is referred by numeral **220'**. The remained engineered cladding layer **220'** may include silicon nitride layer **406** and dielectric layer **404** (associated with FIGS. **5A-5F**); or silicon nitride layer **406** (associated with FIGS. **6A-6H**); or dielectric layer **404** and carbon-rich oxide layer **410** or carbon-rich oxide layer **410** (associated with FIGS. **7A-7F**) according to various embodiments.

[0092] Fabrication may proceed to continue fabrication of multigate device **200**. For example, various contacts can be formed to facilitate operation of the n-type GAA transistors and the p-type GAA transistors. For example, one or more ILD layers, similar to ILD layer **270**, and/or CESL layers can be formed over substrate **202** (in particular, over ILD layer **270** and gate structures **240**). Contacts can then be formed in ILD layer **270** and/or ILD layers disposed over ILD layer **270**. For example, contacts are respectively electrically and/or physically coupled with gate structures **240** and contacts are respectively electrically and/or physically coupled to source/drain regions of the n-type GAA transistors and the p-type GAA transistors (particularly, epitaxial source/drain features **260A**, **260B**). Contacts include a conductive material, such as metal. Metals include aluminum, aluminum alloy (such as aluminum/silicon/copper alloy), copper, copper alloy, titanium, titanium nitride, tantalum, tantalum nitride, tungsten, polysilicon, metal silicide, other suitable metals, or combinations thereof. The metal silicide may include nickel silicide, cobalt silicide, tungsten silicide, tantalum silicide, titanium silicide, platinum silicide, erbium silicide, palladium silicide, or combinations thereof. In some implementations, ILD layers disposed over ILD layer **270** and the contacts (for example, extending through ILD layer **270** and/or the other ILD layers) are a portion of the MLI feature described above.

[0093] Other fabrication processes may be applied to the workpiece **200** and may implemented

before, during, or after the processes described above, such as various processing steps to form an interconnect structure over the GAA transistors from the frontside of substrate **202** to electrically connects various circuit components including the first GAA transistors. The interconnect structure includes metal lines distributed in multiple metal layers (such as 1.sup.st metal layer, 2.sup.nd metal layer, 3.sup.rd metal layer, and etc. from the bottom up to the top metal layer) to provide horizontal routing and contact features (between the substrate and the first metal layer) and via features (between the metal layers) to provide vertical routing. The multigate device **200** also includes other components, such as other conductive features (such as redistribution layer or RDL), passivation layer(s) to provide sealing effect, and/or bonding structures to provide an interface between the multigate device **200** and a circuit board (such as a printed circuit board) to be formed on the interconnect structure.

[0094] In some embodiments, the source/drain features may be formed by other suitable method. For example, the method includes first etching to recess the source/drain regions; a deposition process to deposit one or more sacrificial material layers on sidewalls of the dielectric fins; and epitaxially growth with one or more semiconductor material. The method further includes second etching to remove the deposited sacrificial material layer leaving airgaps between the epitaxial grown source/drain feature and the adjacent dielectric fin. The deposition is designed to form the sacrificial layer with desired shape, such that the final S/D feature has a bar or lollipop like shape. For example, the deposition may include one or more deposition process controlled to have respective deposition characteristics, such as one with conformal deposition (such as atomic layer deposition), another one with more directional deposition (such as bottom-up deposition, or plasma deposition with higher bias power), and another one with more lateral deposition. In some embodiments, the deposited layer is trimmed to achieve the desired shape by a suitable process, such as performing tilted plasma treatment to the upper portions of the deposited dielectric layer, and then another etching process to partially remove the lower portions of the deposited dielectric layer.

[0095] The present disclosure provides multigate device and method making the same for many different embodiments. The method includes forming an engineered cladding layer which includes a semiconductor layer and one or more dielectric layer according to various embodiments. The formation of the engineered cladding layer may include a rapid thermal nitridation (RTN) process, an in-situ carbon (ISC) process and/or other processing steps according to various embodiments. The engineered cladding layer is a composite layer and may include multiple films, such as a semiconductor film and one or more dielectric film formed by RTN, ISC and/or other method. The engineered cladding layer provides paths to selectively remove the second semiconductor layers and release the channels. By implementing the disclosed device structure and the method making the same in various embodiments, some of advantages described below may present. However, it is understood that different embodiments disclosed herein offer different advantages and that no particular advantage is necessarily required in all embodiments. Furthermore, the engineered cladding layer can effectively reduce the diffusion from silicon germanium cladding to the semiconductor fins, and reduce cladding defects, such as cladding footing issue.

[0096] In one example aspect, the present disclosure provides a method of making a semiconductor device. The method includes forming a semiconductor stack on a substrate, wherein the semiconductor stack includes first semiconductor layers of a first semiconductor material and second semiconductor layers of a second semiconductor material alternatively stacked on the substrate; patterning the semiconductor stack and the substrate to form a trench and an active region being adjacent the trench; epitaxially growing a liner of the first semiconductor material on sidewalls of the trench and sidewalls of the active region; forming an isolation feature in the trench; performing a rapid thermal nitridation process, thereby converting the liner into a silicon nitride layer; and forming a cladding layer of the second semiconductor material over the silicon nitride layer.

[0097] In another example aspect, the present disclosure provides a method of making a semiconductor device. The method includes forming a semiconductor stack on a substrate, wherein the semiconductor stack includes first semiconductor layers of a first semiconductor material and second semiconductor layers of a second semiconductor material alternatively stacked on the substrate; patterning the semiconductor stack and the substrate to form a trench and an active region being adjacent the trench; forming an isolation feature in the trench; forming a diffusion blocking layer on sidewalls of the active region; and forming a cladding layer of the second semiconductor material on the diffusion blocking layer.

[0098] In yet another example aspect, the present disclosure provides a method of making a semiconductor device. The method includes forming a semiconductor stack on a substrate, wherein the semiconductor stack includes first semiconductor layers of a first semiconductor material and second semiconductor layers of a second semiconductor material alternatively stacked on the substrate; patterning the semiconductor stack and the substrate to form a trench and an active region being adjacent the trench; epitaxially growing a silicon layer on sidewalls of the trench and sidewalls of the active region, thereby forming a liner; forming an isolation feature in the trench; etching to remove exposed portion of the liner; and epitaxially growing a cladding layer of the second semiconductor material such that the cladding layer is grown selectively on the active region relative to the isolation feature, wherein the cladding layer is in a crystalline structure.

[0099] The foregoing outlines features of several embodiments so that those skilled in the art may better understand the aspects of the present disclosure. Those skilled in the art should appreciate that they may readily use the present disclosure as a basis for designing or modifying other processes and structures for carrying out the same purposes and/or achieving the same advantages of the embodiments introduced herein. Those skilled in the art should also realize that such equivalent constructions do not depart from the spirit and scope of the present disclosure, and that they may make various changes, substitutions, and alterations herein without departing from the spirit and scope of the present disclosure.

Claims

1. A method, comprising: forming a semiconductor stack on a substrate, wherein the semiconductor stack includes first semiconductor layers of a first semiconductor material and second semiconductor layers of a second semiconductor material alternatively stacked on the substrate; patterning the semiconductor stack and the substrate to form a trench and an active region being adjacent the trench; epitaxially growing a liner of a third semiconductor material on sidewalls of the active region; forming an isolation feature over the liner within the trench; performing a rapid thermal nitridation process, thereby converting a portion of the liner into a semiconductor nitride layer; and forming a cladding layer of a fourth semiconductor material over the semiconductor nitride layer.
2. The method of claim 1, further comprising forming a dielectric layer on the liner before the performing of the rapid thermal nitridation process.
3. The method of claim 2, further comprising removing the dielectric layer after the performing of the rapid thermal nitridation process and before the forming of the cladding layer.
4. The method of claim 3, wherein the dielectric layer includes silicon oxide; and the removing the dielectric layer further includes performing a chemical oxide removal (COR) process, and performing a post heating treatment (PHT).
5. The method of claim 4, wherein the COR process includes applying NH_3 and HF at a first temperature; and the PHT includes applying an annealing process at a second temperature greater than the first temperature.
6. The method of claim 1, wherein the first semiconductor material is silicon; and the second semiconductor material is silicon germanium.

7. The method of claim 1, wherein the third semiconductor material is silicon; the fourth semiconductor material is silicon germanium; and the semiconductor nitride layer is a silicon nitride layer.
8. The method of claim 1, wherein the forming of the cladding layer includes depositing the cladding layer on the active region such that the cladding layer is separated from the semiconductor stack by the semiconductor nitride layer; and performing an anisotropic etching process to the cladding layer.
9. The method of claim 8, wherein the depositing the cladding layer on the active region includes depositing the cladding layer with a first portion on the semiconductor nitride layer and a second portion on the liner; the first portion of the cladding layer has an amorphous structure; and the second portion of the cladding layer has a crystalline structure.
10. The method of claim 1, further comprising: forming a dummy gate stack on the active region; forming source/drain (S/D) features on the active region and connecting to the first semiconductor layers; removing the dummy gate stacks, resulting in a gate trench in an interlayer dielectric (ILD) layer; performing an etching process in the gate trench to selectively remove the cladding layer and the second semiconductor layers; and forming a gate stack in the gate trench, the gate stack wrapping around each of the first semiconductor layers.
11. A method, comprising: forming a semiconductor stack on a substrate, wherein the semiconductor stack includes first semiconductor layers of a first semiconductor material and second semiconductor layers of a second semiconductor material alternatively stacked on the substrate; patterning the semiconductor stack and the substrate to form a trench and an active region being adjacent the trench; epitaxially growing a liner of the first semiconductor material on sidewalls of the active region; forming an isolation feature over the liner within the trench; and forming a diffusion blocking layer on the sidewalls of the active region and on the isolation feature.
12. The method of claim 11, further comprising forming a cladding layer of the second semiconductor material on the diffusion blocking layer.
13. The method of claim 12, further comprising: forming a dummy gate stack on the cladding layer; forming source/drain (S/D) features on the active region and connecting to the first semiconductor layers; removing the dummy gate stacks, resulting in a gate trench in an interlayer dielectric (ILD) layer; performing an etching process in the gate trench to selectively remove the cladding layer and the second semiconductor layers; and forming a gate stack in the gate trench, the gate stack wrapping around each of the first semiconductor layers.
14. The method of claim 12, wherein the diffusion blocking layer includes a silicon oxide layer; the first semiconductor material is silicon; and the second semiconductor material is silicon germanium.
15. The method of claim 14, wherein the forming of the diffusion blocking layer further includes performing an in-situ carbon (ISC) process to form a carbon-rich dielectric layer on the silicon oxide layer.
16. The method of claim 15, wherein the performing an ISC process includes applying a precursor containing ##STR00001##
17. The method of claim 14, further comprising performing a rapid thermal nitridation (RTN) process, thereby converting the liner into a silicon nitride layer.
18. The method of claim 17, further comprising removing the silicon oxide layer after the performing of the rapid thermal nitridation process and before the forming of the cladding layer, wherein the removing the dielectric layer includes applying NH₃ and HF at a first temperature; and applying a post heating treatment at a second temperature being greater than the first temperature.
19. A method, comprising: forming a semiconductor stack on a substrate, wherein the semiconductor stack includes first semiconductor layers of a first semiconductor material and second semiconductor layers of a second semiconductor material alternatively stacked on the substrate; patterning the semiconductor stack and the substrate to form a trench and an active

region being adjacent the trench; epitaxially growing a silicon layer on sidewalls of the active region, thereby forming a liner; forming an isolation feature on the isolation feature in the trench; and etching to remove exposed portion of the liner above the isolation feature.

20. The method of claim 19, further comprising epitaxially growing a cladding layer of the second semiconductor material such that the cladding layer is selectively grown on the active region relative to the isolation feature, wherein the cladding layer is in a crystalline structure and is free from surfaces of the isolation feature.
