

Downstream etching of silicon nitride using continuous-wave and pulsed remote plasma sources sustained in Ar/NF₃/O₂ mixtures

Shuo Huang^{a)}

Department of Electrical Engineering and Computer Science, University of Michigan, 1301 Beal Ave., Ann Arbor, Michigan 48109-2122

Vladimir Volynets^{b)}

Samsung Electronics Co., Ltd., 129 Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 443-742, South Korea

James R. Hamilton^{c)}

Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom

Sang Ki Nam,^{d)} In-Cheol Song,^{e)} and Siqing Lu^{f)}

Samsung Electronics Co., Ltd., 129 Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 443-742, South Korea

Jonathan Tennyson^{g)}

Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom

Mark J. Kushner^{h)}

Department of Electrical Engineering and Computer Science, University of Michigan, 1301 Beal Ave., Ann Arbor, Michigan 48109-2122

(Received 15 December 2017; accepted 5 February 2018; published 26 February 2018)

Remote plasma sources (RPSs) are being investigated to produce fluxes of radicals for low damage material processing. In this computational investigation, the properties of a RPS etching system are discussed where an Ar/NF₃/O₂ gas mixture is flowed through an inductively coupled plasma source into a downstream chamber containing a silicon nitride coated wafer. The plasma is largely confined in the RPS due to the highly attaching NF_x (x = 1–3) and an isolating showerhead although a weak ion-ion plasma maintained by [NO⁺] ≈ [F⁻] leaks into the downstream chamber. The etching of silicon nitride proceeds through iterative removal of Si and N subsites by isotropic thermal neutrals. When the fluxes to the wafer are rich in fluorine radicals, the etch rate is limited by the availability of NO molecules and N atoms which remove N subsites. As power deposition increases with continuous-wave excitation, the etch rate increases almost linearly with the increasing fluxes of NO and N atoms, as production of NO through endothermic reactions is aided by increasing gas temperature. Production of N atoms through electron impact dissociation of NO and NF_x is aided by the increasing electron density. Similar trends occur when increasing the duty cycle during pulsed excitation. Addition of a plenum between the RPS and the downstream chamber aids in lateral diffusion of radicals before passing through the final showerhead and improves the uniformity of etching. *Published by the AVS.* <https://doi.org/10.1116/1.5019673>

I. INTRODUCTION

Remote plasma sources (RPSs) are being developed for isotropic etching and thin film deposition in microelectronic fabrication.^{1,2} Due to long flow distances, grids, or other discriminating barriers between the RPS and the substrate, the incident fluxes onto the wafer consist primarily of neutral radicals, and so, damage by charging, energetic ion bombardment, and ultraviolet (UV) or vacuum-ultraviolet (VUV) radiation can be minimized.³ Due to the absence of energetic ions that can produce physical sputtering, RPS provides a mechanism for high etch selectivity based only on

chemical mechanisms. For example, high etch selectivity of Si₃N₄ over SiO₂ can be achieved by increasing the flow rate of O₂ in a RPS sustained in NF₃/O₂ (Ref. 4) or by increasing the flow rate of N₂ through a RPS sustained in CF₄/O₂/N₂.⁵ This selectivity results from enhanced utilization of O and NO radicals in the etching of Si₃N₄ compared to SiO₂. RPS has also been implemented for processing materials such as low-k (porous SiOCH) interlayer dielectrics⁶ and high-k (Al₂O₃ and SiC) gate dielectrics.⁷ These remote processes minimize surface roughness and mixing typically produced by energetic ion bombardment. Recently, three-dimensional structures such as stacked planar NAND (Ref. 8) and vertical NAND (Ref. 9) have been proposed to increase device density for flash memories. In these structures, RPS can be used to selectively etch the oxide and nitride in the horizontal direction between layers in the stack.

Gas temperature is, in principle, a controllable process variable in RPS systems that can be used to produce desired reactants. The gas temperature, T_g , can be elevated in the RPS while near ambient in the downstream etch chamber.

^{a)}Electronic mail: shuoh@umich.edu

^{b)}Electronic mail: vladimir.volynets@samsung.com

^{c)}Electronic mail: james.hamilton@ucl.ac.uk

^{d)}Electronic mail: sangki.j.nam@samsung.com

^{e)}Electronic mail: ic13.song@samsung.com

^{f)}Electronic mail: siqing.lu@samsung.com

^{g)}Electronic mail: j.tennyson@ucl.ac.uk

^{h)}Author to whom correspondence should be addressed; electronic mail: mjkuhs@umich.edu

Endothermic reactions are then isolated to the RPS, uniquely producing reactants that are not generated downstream. For example, flowing an NF_3/O_2 mixture through a RPS with high power deposition generating T_g exceeding 400–500 K will produce N_xO_y species in the source through endothermic reactions. The production of N_xO_y will cease downstream where T_g cools.¹⁰ RPS also provides the possibility of selectively generating species by mixing of gases. For example, by flowing NF_3 gas into the downstream of a RPS sustained in H_2 (Ref. 11) or $\text{H}_2/\text{H}_2\text{O}$,¹² hot H atoms produced in the RPS react with NF_3 to form HF through exothermic reactions, an isotropic etchant of SiO_2 . Due to the low gas temperature and the absence of potentially damaging plasma in the downstream chamber, RPS serves as a mild and conformal method for fabricating atomic scale temperature-sensitive materials such as TiO_2 on Si substrates¹³ and surface oxidation in the atomic layer etching of MoS_2 .¹⁴

NF_3 is frequently used in RPS for the ease with which F atoms, the main etchants of silicon-containing materials such as SiO_2 , SiC, and Si_3N_4 , are produced by dissociative attachment and dissociative excitation. Adding other gases to NF_3 increases the variety of reactive species that can be produced. For example, the use of NF_3/N_2 mixtures in RPS increases the etch rate of Si through enhancing the decomposition of NF_3 by N radicals and the formation of F radicals.¹⁵ The use of NF_3/O_2 mixtures increases the etch rate of Si_3N_4 by producing NO molecules which aid in the removal of N atoms from the surface.⁴ The use of NF_3/NH_3 mixtures increases the etch rate of SiO_2 by forming a condensed layer on the surface containing ionic active species such as NH_4 and F which preferentially attack the Si–O bond.¹⁶ For certain applications, it may be desirable to separately optimize, for example, F and N_xO_y fluxes to optimize the etch rates of Si, SiO_2 , Si_3N_4 , and other materials. This separate optimization could, in principle, be performed using pulsed power or pulsed gas sources.

Chemical downstream systems for the etching of polysilicon and SiO_2 using NF_3/O_2 mixtures were modeled by Meeks *et al.*¹⁷ and Vosen *et al.*¹⁸ In these efforts, the RPS, the transport tube, and the downstream etch chamber in tandem were modeled using a zero-dimensional well mixed reactor model, a one-dimensional plug flow model, and a two-dimensional axisymmetric reacting-flow model, respectively. The etch rate of polysilicon was found to be transport limited, whereas the etch rate of SiO_2 was found to be surface-kinetics limited, depending more on pressure and less on the flow rate than the etching of polysilicon.

Etch rates of Si_3N_4 , SiO_2 , and polysilicon and radical densities in the RPS measured by optical emission spectroscopy and global modeling were recently reported by Barsukov *et al.* for a NF_3/O_2 remote plasma.¹⁹ They described an etching mechanism for Si_3N_4 , including the role of NO in the etching of Si_3N_4 by atomic fluorine. Quantum chemistry modeling of the surface kinetics showed that NO reacts with the F–N bond to form N_2O , and the resulting fluorine migrates from the F–N bond to the neighboring silicon atom. By NO enhancing the rate of fluorine

migration on the Si_3N_4 surface from the nitrogen atom to the silicon atom, the etch rate of Si_3N_4 is increased.

In this paper, the kinetics occurring in a RPS sustained in $\text{Ar}/\text{NF}_3/\text{O}_2$ and the downstream etching of Si_3N_4 are discussed using results from a two-dimensional model. The simulation includes the plasma kinetics and flow dynamics in the RPS, the transport channel, and the downstream chamber. A surface site balance model is used to represent etching at the surface of the Si_3N_4 covered wafer. Both continuous-wave (CW) and pulsed systems are addressed. The electron impact cross-sections for NF_2 and NF were produced using *ab initio* computational techniques based on the molecular R-matrix method.

We found that plasmas in this system are largely confined in the source region due to isolating showerheads and the highly attaching nature of the gas mixture. The main etchants of silicon nitride, F and N containing neutrals, are produced through electron impact reactions and endothermic heavy particle reactions in the RPS. These species flow downstream with densities that rebound due to gas cooling from thermal conduction to the walls. Silicon nitride is etched by iteratively removing Si subsites by F atoms and removing N sites by NO and N. With fluorine rich fluxes to the wafer, the etch rate can be increased by using operating conditions that increase the gas temperature in the source [e.g., increasing the duty cycle (DC) during pulsed operation] to increase the formation of NO and N. Adding a plenum between the RPS and the downstream etch chamber increases the uniformity of radical fluxes to the wafer surface and thus increases the uniformity of the silicon nitride etch rate.

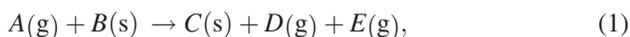
A description of the computational model is in Sec. II. The surface reaction mechanism for the etching of silicon nitride by a RPS sustained in $\text{Ar}/\text{NF}_3/\text{O}_2$ mixtures is described in Sec. III. The scaling of plasma and etch properties with operating conditions for RPS sustained by CW and pulsed power is discussed in Secs. IV and V. Concluding remarks are presented in Sec. VI.

II. DESCRIPTION OF THE MODEL

Reactor scale simulation of the downstream etch system was performed using the two-dimensional Hybrid Plasma Equipment Model (HPEM).²⁰ The HPEM is a kinetic-fluid hydrodynamics code which combines separate modules that address different physical phenomena in an iterative manner. In this investigation, the major modules used are the Electromagnetics Module (EMM), the Electron Energy Transport Module (EETM), the Fluid Kinetics Poisson module (FKPM), and the Surface Kinetics Module (SKM). Inductively coupled electromagnetic fields are produced by the EMM using a frequency domain solution of Maxwell's equations. The densities of all charged and neutral species and the electrostatic potentials are obtained from the FKPM. Continuity, momentum, and energy equations for all species are solved coincident with Poisson's equation for the electric potential. Charge densities on the surfaces of the dielectrics are computed from the fluxes of electrons and ions from the bulk

plasma and secondary electrons leaving the surfaces and coming from other locations collected by those surfaces. The electron energy equation is implicitly integrated in time to provide the electron temperature, impact rate coefficients, and transport coefficients. These coefficients as a function of average electron energy are provided by solutions of Boltzmann's equation for the electron energy distribution. The trajectories of secondary electrons are advanced using Monte Carlo techniques in the EETM.

The etching process and the coupling between the bulk and surface processes are addressed by the SKM in which a surface site balance model is executed. The model consists of rate equations for the fractional coverage of surface resident species based on reactions with fluxes from the gas phase or between surface sites. Based on these surface coverages, the reactive sticking coefficients for gas phase species on surfaces are updated.²¹ The fluxes of gas phase species to the surface used in the SKM are provided by the results of the FKPM. The general form of a plasma-surface reaction is



where g denotes a gas phase species and s denotes a surface site. The rate for the i th reaction in the surface reaction mechanism between gas phase species A and surface site B on material m is

$$R_{im} = k_i \Phi_{Am} \theta_{Bm}, \quad (2)$$

where k_i is the reaction probability of the i th reaction, Φ_{Am} is the incident flux of gas phase species A on material m , and θ_{Bm} is the fractional occupancy of the surface sites by species B . The rates of evolution of the coverages of surface sites are obtained by summing up the rates of reactions generating or consuming the sites. The steady state coverages of all surface sites are obtained by integrating the coupled rate equations for all surface sites using a third-order Runge–Kutta technique. The total time length to integrate the rate equations for the surface reactions is a few seconds, which is long enough for the results to converge. The reaction probability for an incident gas phase species is then the sum of fractional losses by all reactions removing the species.

Several modifications were made to the HPEM to increase the numerical efficiency and better represent the regions of the RPS which have significantly different mole fractions of reactants. The HPEM uses a rectilinear structured mesh with computational loops that are nested by, for example, radial and axial mesh points. OpenMP parallelization directives are employed on the outer loop for computational efficiency. An outer-mesh covers the entire computational domain, which includes nonplasma points (e.g., dielectrics, surrounding air, and electrical ground planes). Within the outer-mesh, a sub-mesh covers a region which captures the plasma. Due to the rectilinear structure, even within the submesh computational domain, only a subset of mesh points may represent the plasma, with the remainder representing solid materials (e.g., flow tubes, substrates, and electrodes) or surrounding air. In order to

maximize numerical pipeline, prefetch, and parallelization efficiencies, all plasma transport equations are solved at all points in the submesh. The pipeline efficiency resulting from looping over the structured mesh provides higher net computational speed even if some of the mesh points are nonplasma. A stencil is used to identify plasma points and zero out calculations at nonplasma points. If the fraction of the submesh points that are not plasma is sufficiently small, then the pipeline and prefetch efficiencies win out over the additional work to compute at nonplasma points, and the net computing speed is higher.

For those conditions where a large fraction of the sub-mesh is not plasma, the just described technique is not computationally efficient. For those cases, a node-and-neighbor technique is used to loop over only plasma points. A list of plasma mesh points and their nearest neighbors is constructed, and OpenMP loops are performed over the plasma node-list. Although this is not as efficient with respect to prefetch opportunities, the end result is faster over all computations if at least $\approx 1/3$ of the submesh is not plasma. To deploy this technique, several node-lists are required—for plasma points at which densities, temperatures, and potentials are computed, for radial and for axial momenta which are computed at the half-points between plasma points, and for nonplasma points bounding plasma points.

When using the electron energy equation, rate coefficients for collisional processes (e.g., electron impact ionization) are required as a function of electron temperature, T_e , or average electron energy, $\langle \epsilon_e \rangle$. This is accomplished by solving Boltzmann's equation for the electron energy distribution over a wide range of E/N (electric field/gas number density) values and constructing a look-up table for rate coefficients as a function of T_e . This table is periodically updated as the mole fractions of species change due to reactions. In a RPS, the mole fractions of gases may greatly vary as a function of position. For these conditions, the numerical mesh is divided into different regions or zones which individually cover volumes of the plasma having significantly different mole fractions. Separate lookup tables for rate coefficients are generated for each region to address these spatially dependent differences in mole fractions. Although there is computational expense associated with generating and updating multiple look-up tables, the resulting rate coefficients are better representations than using volume averaged rate coefficients over all plasma regions.

III. SURFACE REACTION MECHANISM

The goal of many RPSs is to confine ions, energetic particles, and UV/VUV radiation to the remote source, resulting in the wafer being dominantly exposed to neutral radical fluxes. This investigation addresses the etching of Si_3N_4 by the neutral fluxes produced downstream of a RPS sustained in Ar/ NF_3/O_2 mixtures. The surface reaction mechanism for our surface site balance model is listed in Table I. The notation of SiN is used for silicon nitride to simplify the reaction mechanism and mapping of species to surface sites. Although SiN is not strictly the same stoichiometry as Si_3N_4 , this simpler model does provide insight into scaling relationships.

TABLE I. Surface reaction mechanism for etching of silicon nitride by a remote plasma source sustained in Ar/NF₃/O₂ mixtures.

Gas phase species					
F(g)	N ₂ (g)	N(g)	O ₂ (g)	O(g)	NO(g)
N ₂ O(g)	NO ₂ (g)	SiF ₄ (g)			
Surface sites ^a					
SiN(s)	SiNF(s)	SiNF ₂ (s)	SiNF ₃ (s)	N(s)	O(s)
Reactions ^b		Probability ^{c, d}		References	Note
Removal of Si subsites					
SiN(s) + F(g) → SiNF(s)		0.1			c
SiNF(s) + F(g) → SiNF ₂ (s)		0.1			c
SiNF ₂ (s) + F(g) → SiNF ₃ (s)		0.1			c
SiNF ₃ (s) + F(g) → SiF ₄ (g) + N(s)		0.1			c
Removal of N subsites					
N(s) + NO(g) → N ₂ O(g) + SiN(s)		0.1		22	c
N(s) + N(g) → N ₂ (g) + SiN(s)		0.1			c
Surface oxidation to form O sites					
N(s) + NO(g) → N ₂ (g) + O(s)		0.1		22	c
Removal of O sites					
O(s) + NO(g) → NO ₂ (g) + SiN(s)		0.1		30	c
O(s) + O(g) → O ₂ (g) + SiN(s)		0.1			c
O(s) + O(s) → O ₂ (g) + SiN(s)		10 ¹⁸			d

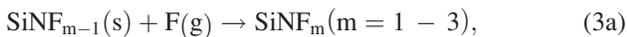
^aThe notation of SiN is used for silicon nitride to simplify the reaction mechanism and primarily achieve balance of surface sites.

^bAll gas phase species have units of flux (cm⁻²s⁻¹). All surface species have units of fractional coverage (unitless).

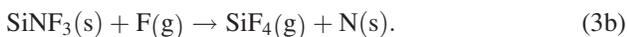
^cProbability for reactions between gas phase species and surface sites is unitless.

^dProbability for reactions between surface sites and surface sites is cm⁻²s⁻¹. Probability is obtained by $R \times T^2$, where R is the surface reaction rate coefficient and T is the surface site density. In this study, $R = 10^{-12}$ cm²s⁻¹ and $T = 1 \times 10^{15}$ cm⁻².

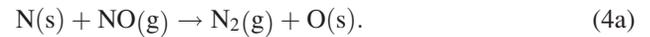
The initial surface composition of the wafer consists of pristine SiN sites containing Si and N subsites. The pristine SiN sites are consecutively passivated by F atoms through



and the Si subsites are etched by the formation of volatile SiF₄ through



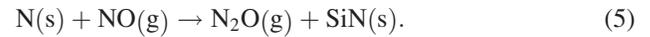
The notation N(s) represents the exposed N subsites in SiN(s) but still bonded to the substrate, whose removal is less straightforward. There have been several proposals for removing N(s) in similar gas mixtures. One proposal is based on the results from x-ray photoelectron spectroscopy measurements in the etching of silicon nitride by a RPS sustained in the NF₃/O₂ mixture.²² The proposed removal process consists of two channels. The first is initiated by reaction with NO(g) which removes N surface sites by the formation of N₂ with the adsorption of O onto the surface (i.e., surface oxidation) through



The O surface sites subsequently desorb through the formation of NO₂ and O₂, and the underlying SiN sites are exposed,



In the second channel, the N sites are etched with the formation of N₂O through



Alternately, gas phase N atoms can remove N sites which then expose the underlying SiN through



The etching of SiN sites proceeds by iteratively etching the Si and N subsites. As these surface reactions are all thermally driven and occur in the absence of directional ion bombardment, the etching mechanism described in Table I is isotropic.

IV. REMOTE PLASMA SOURCE SUSTAINED BY CW INDUCTIVELY COUPLED POWER

A schematic of the downstream etch system as implemented in the HPEM with one remote inductively coupled source sustained in Ar/NF₃/O₂ is shown in Fig. 1. The species in the model are Ar, Ar(1s₁), Ar(1s₂), Ar(1s₃), Ar(1s₄), Ar(4p), Ar(4d), Ar₂(³Σ_u⁺), Ar⁺, Ar₂⁺, NF₃, NF₂, NF,

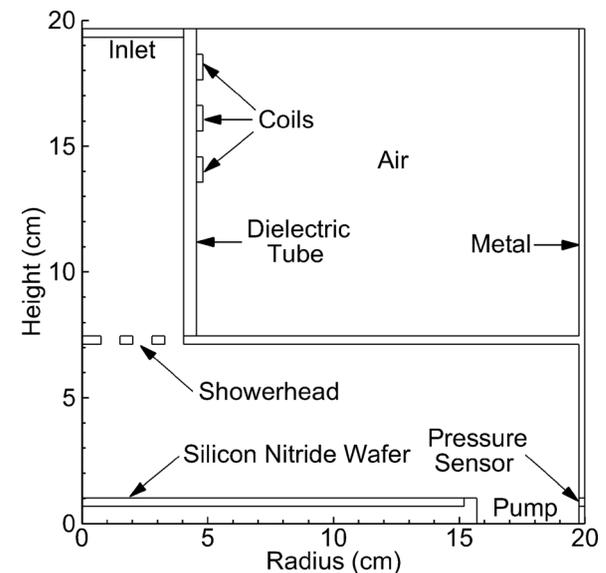


Fig. 1. Schematic of a downstream etch system consisting of a remote plasma source and a downstream reactor chamber with metal showerheads placed in-between. The RPS is driven by inductively coupled power. The silicon nitride wafer is placed on the substrate in the downstream chamber with no bias imposed on it.

NF_3^+ , NF_2^+ , NF^+ , N_2 , $\text{N}_2(\text{v})$, $\text{N}_2(\text{A } ^3\Sigma_u^+)$, $\text{N}_2(\text{B } ^3\Pi_g, \text{higher})$, N , $\text{N}(^2\text{D})$, N_2^+ , N^+ , F_2 , $\text{F}_2(\text{I } ^1\Sigma_u^+)$, F , $\text{F}(^3\text{S})$, F_2^+ , F^+ , F^- , O_2 , $\text{O}_2(\text{v})$, $\text{O}_2(\text{a } ^1\Delta_g)$, $\text{O}_2(\text{b } ^1\Sigma_g^+)$, O , $\text{O}(^1\text{D})$, $\text{O}(^1\text{S})$, O_3 , O_2^+ , O^+ , O_2^- , O^- , O_3^- , FO , FNO , NO , N_2O , NO_2 , NO^+ , N_2O^+ , and electrons. The reaction mechanism is the same as described in Ref. 23. The electron impact cross-sections for NF_2 and NF were produced using *ab initio* computational techniques based on the molecular R-matrix method.^{24,25} The $\text{Ar}/\text{NF}_3/\text{O}_2$ mixture was pumped in through the inlet at the top of the RPS, dissociated in the RPS by the inductively coupled power to produce reactive species for downstream etching. The radius of the RPS is 4 cm. The wafer was placed 12 cm downstream from the RPS, with metal showerheads placed in-between. The spacing of the showerheads is 0.75 cm, which even though is several times of the local Debye length (≈ 0.1 cm) still isolates the plasma from reaching the downstream chamber. The silicon nitride wafer was placed on the substrate with no applied bias. A pressure sensor was placed near the pump port, and the pumping rate was adjusted to maintain the gas pressure at 50 mTorr at the location of the sensor. For the base case, the inlet gas mixture is $\text{Ar}/\text{NF}_3/\text{O}_2 = 90/5/5$ with a flow rate of 1000 sccm. The total power delivered from the inductively coupled plasma (ICP) coils is 300 W. The intent of this work is to investigate reaction mechanisms and not necessarily to optimize the uniformity of processing.

Electrons are confined in the RPS region due to the highly attaching nature of the gas mixture and the limiting nature of the showerhead. In the steady state, the electron density is $3.7 \times 10^{11} \text{ cm}^{-3}$ near the ICP coils, where the electron temperature peaks at up to 4.5 eV as shown in Fig. 2. NF_3 and its dissociation products, NF_2 and NF , are all thermal attaching molecules as the electron affinity of F (3.4 eV) is larger than the binding energy of NF_3 ($D_0 = 2.4$ eV),²⁶ NF_2 ($D_0 = 2.9$ eV)²⁶ and NF ($D_0 = 2.8$ eV),²⁷ and so, a high electron temperature is required to provide ionization to balance the high rate of attachment. The ICP power is deposited within the skin depth of the radio frequency field ($\approx 1\text{--}2$ cm), and only those electrons located in the skin depth are initially heated. Electrons outside the skin depth are heated by convection and thermal conduction. Combined with collisional losses, T_e decreases to 3.6 eV at the confining grid of the RPS. The net electron impact ionization source, S_e , peaks near the coils at $6.2 \times 10^{16} \text{ cm}^{-3} \text{ s}^{-1}$ as shown in Fig. 2(c). The dependence on T_e of both electron attachment (decreasing from low to high T_e) and ionization (increasing from low to high T_e) is quite steep, resulting in a clear demarcation between ionization dominated with net gain by electron impact (high T_e , upstream) and attachment dominated with net loss by electron impact (low T_e , downstream). This transition occurs at ≈ 4.0 eV, which produces a sharp boundary between the electron impact source function between positive and negative, as shown in Fig. 2(c). This line of demarcation occurs at higher T_e at the top of the RPS zone than at the bottom of the RPS zone, due to the dissociation of NF_3 flowing downstream which reduces the rate of attachment.

The electrons are largely confined to the upstream RPS zone. This occurs for at least two reasons—the first is the

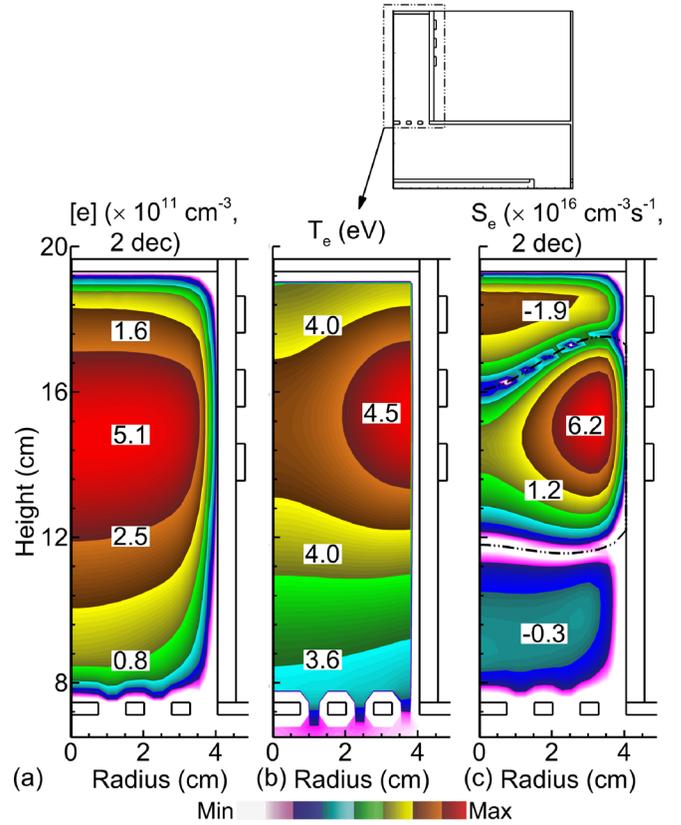


Fig. 2. (Color online) Time averaged (a) electron density, (b) electron temperature, and (c) electron impact ionization source in a RPS driven by inductively coupled power. Operating conditions: $\text{Ar}/\text{NF}_3/\text{O}_2 = 90/5/5$, 50 mTorr, 1000 sccm, 300 W, and 10 MHz.

physical confinement by the grids and the second is the rapid transition of the plasma to an ion-ion system. The rate of attachment is sufficiently rapid as T_e decreases through the grids that the electrons are essentially totally depleted by attachment. Due to the large grid spacing, there is a small leakage of plasma through the grids; this leakage is essentially an ion-ion plasma. The densities of the positive and negative ions are 10^{11} cm^{-3} in the RPS, while leakage to the downstream chamber produces a density of 10^8 cm^{-3} as shown in Fig. 3. The dominant positive ions in the RPS region are Ar^+ (60%) and O_2^+ (25%) as shown in Figs. 3(c) and 3(d) due to the large mole fraction of their parent gases. The dominant negative ion is F^- as shown in Fig. 3(f) due to thermal attaching NF_x ($x = 1 - 3$) and the highest electron affinity of F among all the neutrals in the system. The fractions of NF_3^+ , NF_2^+ , and NF^+ are less than 3% due to high rates of dissociative attachment which decompose NF_x , and moderate ionization potentials of their parent gases (12.9 eV for NF_3 , 11.6 eV for NF_2 , and 12.1 eV for NF).

For electronegative, molecular gas mixtures, the electron loss is dominated by attachment and dissociative recombination and volumetric losses which contribute to the confinement of the plasma in the remote source. For electropositive gases, weakly electronegative gas mixtures, and mixtures dominated by atomic ions, the dominant mechanism for electron loss is diffusion and recombination on surfaces. As this transition occurs between rapid volumetric electron loss

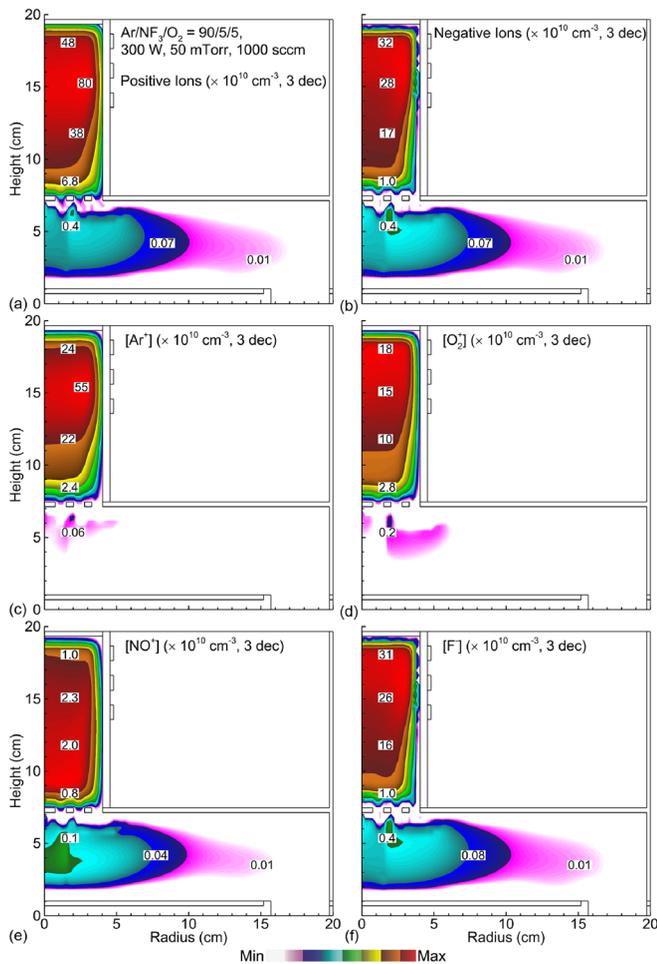


Fig. 3. (Color online) Time averaged densities of (a) all positive ions, (b) all negative ions, (c) Ar^+ , (d) O_2^+ , (e) NO^+ , and (f) F^- in a downstream etch system with a RPS driven by inductively coupled power. Operating conditions: $\text{Ar}/\text{NF}_3/\text{O}_2 = 90/5/5$, 50 mTorr, 1000 sccm, 300 W, and 10 MHz.

and slow diffusional wall losses, isolation of the plasma from the downstream chamber requires longer flow distances to enable time for the wall recombination to occur. Another strategy is to use openings in the showerhead which are smaller than the Debye length, to prevent the plasma from being conformal to the openings.

The weak ion-ion plasma downstream of approximately 10^8 cm^{-3} is maintained by $[\text{NO}^+] \approx [\text{F}^-]$ as shown in Figs. 3(e) and 3(f). There is essentially no electron impact ionization occurring downstream. However, there are production and mixing of positive ions through Penning ionization and charge exchange. As the ionization potential of NO (9.3 eV) is the lowest among all the important neutral species, charge exchange and Penning ionization predominantly favor the formation of NO^+ ions. The few electrons that may be produced by Penning ionization are quickly consumed by dissociative attachment to NF_x . Since the density of ions in the downstream chamber is quite low and the mobilities of F^- and NO^+ are commensurate, there is little ambipolar enhancement of the diffusion rates for the ions and the majority of positive and negative ions are lost through ion-ion neutralization. As a result, there is little accumulated charge on surfaces and the potential damage to the wafer by charging is small.

The gas temperature reaches as high as 1080 K in the RPS, shown in Fig. 4(a), due largely to heating through dissociative attachment and dissociative excitation by the Franck-Condon effect. The contribution to gas heating by symmetric charge exchange (e.g., Ar^+ with Ar) in the ambipolar electric fields of the source reaction is small. In the downstream chamber, the electron density is small and there is essentially no electron impact dissociation that might additionally heat the gas through Franck-Condon processes. The exothermic and endothermic heavy particle reactions result in net heating in the downstream chamber as most of the endothermic reactions are significantly inhibited as gas

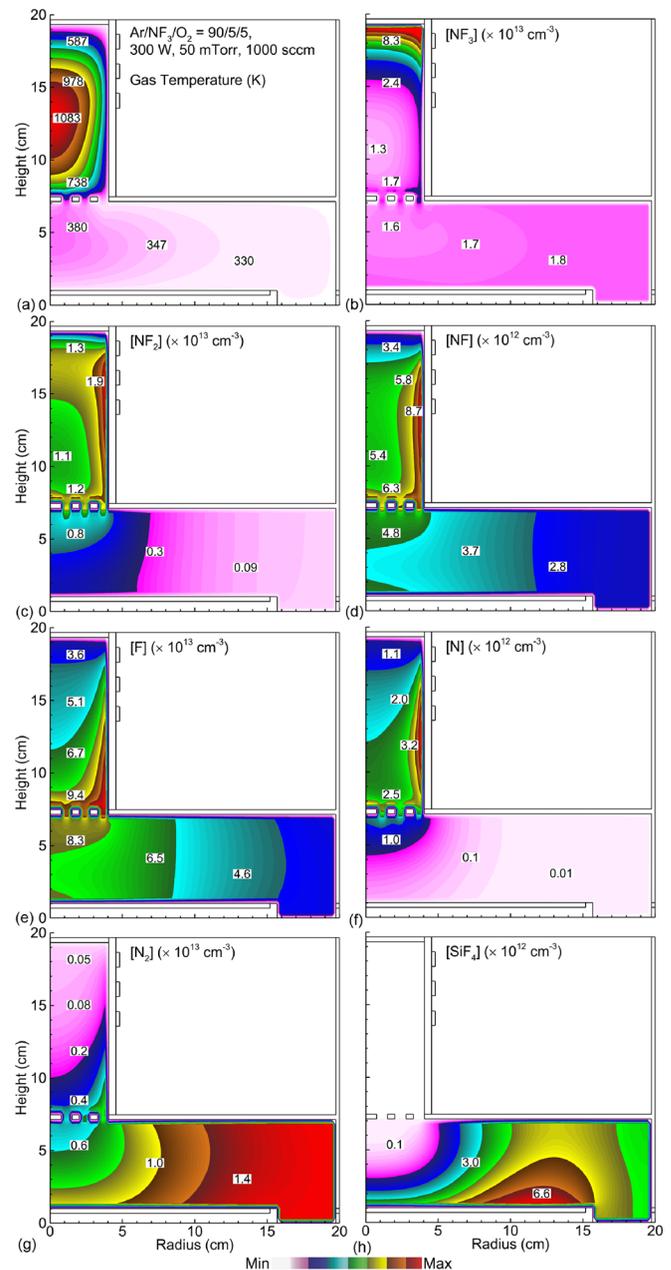


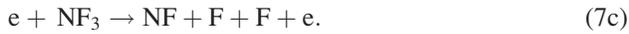
Fig. 4. (Color online) Time averaged (a) gas temperature and densities of (b) NF_3 , (c) NF_2 , (d) NF , (e) F , (f) N , (g) N_2 , and (h) SiF_4 in a downstream etch system with a RPS driven by inductively coupled power. Operating conditions: $\text{Ar}/\text{NF}_3/\text{O}_2 = 90/5/5$, 50 mTorr, 1000 sccm, 300 W, and 10 MHz.

temperature falls below 400 K. With the addition of thermal conduction and in spite of the allowance for slip and temperature jump at surfaces, the gas temperature decreases to nearly ambient near the wafer. Since the system is essentially isobaric, the densities of ambient neutrals (e.g., NF_3 and O_2) rebound with the decrease in gas temperature, as shown in Figs. 4(b) and 5(a).

The densities of NF_3 and its dissociation products, NF_2 , NF , N and F , are shown in Fig. 4. In the RPS region, NF_3 is dissociated to form NF_2 , NF , and F through dissociative electron attachment



and dissociative excitation



The density of NF_3 decreases from 8.3 to $1.7 \times 10^{13} \text{ cm}^{-3}$ with the flow, reaching a fractional dissociation of 72% at the end of the RPS region. The density of F atoms increases from the gas inlet to a maximum of $9.4 \times 10^{13} \text{ cm}^{-3}$ at the

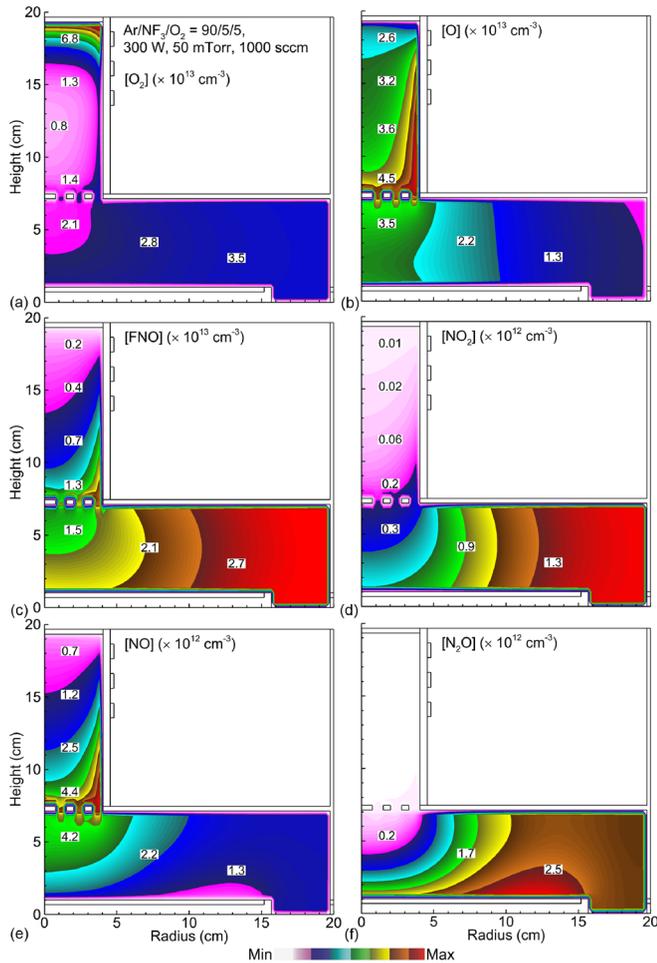


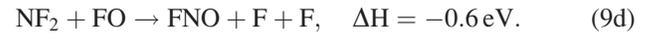
Fig. 5. (Color online) Time averaged densities of (a) O_2 , (b) O , (c) FNO , (d) NO_2 , (e) NO , and (f) N_2O in a downstream etch system with a RPS driven by inductively coupled power. Operating conditions: $\text{Ar}/\text{NF}_3/\text{O}_2 = 90/5/5$, 50 mTorr, 1000 sccm, 300 W, and 10 MHz.

end of the RPS region, becoming the dominant radical. In the downstream chamber where there is essentially no electron impact dissociation, the fractional dissociation of NF_3 slightly increases to 73% at the wafer surface as NF_3 is decomposed by radicals through exothermic reactions



where ΔH is the change in the enthalpy of the reaction (negative ΔH indicates an exothermic reaction). However, the density of NF_3 slightly increases from $1.6 \times 10^{13} \text{ cm}^{-3}$ near the showerheads to $1.8 \times 10^{13} \text{ cm}^{-3}$ at the wafer surface. This rebound in density is caused by the gas temperature decreasing from nearly 1080 K in the RPS source and 740 K at the grids to 330 K at the wafer, as shown in Fig. 4(a). The decrease in gas temperature is due dominantly to thermal conduction to the walls in the downstream chamber.

The densities of NF_2 , NF , and N show similar trends, being formed in the RPS through electron impact reactions and reaching densities of 1.2×10^{13} , 6.3×10^{12} , and $2.5 \times 10^{12} \text{ cm}^{-3}$ at the end of the RPS region, respectively. The densities of NF_2 , NF , and N peak at 1.9×10^{13} , 8.7×10^{12} , and $3.2 \times 10^{12} \text{ cm}^{-3}$ near the coils where the electron impact ionization source and the electron temperature are the maximum and the electron impact dissociation reactions are favored. After entering the downstream chamber, the density of NF_2 decreases to $2.0 \times 10^{12} \text{ cm}^{-3}$ at the wafer surface due to consumption through exothermic reactions



The density of NF also decreases to $3.7 \times 10^{12} \text{ cm}^{-3}$ at the wafer surface due to the exothermic reactions in Eq. (8c) and



The density of N atoms decreases to $1.0 \times 10^{11} \text{ cm}^{-3}$ at the wafer surface in part due to the surface reaction in Eq. (6) [abstraction of $\text{N}(\text{s})$ from the surface] and in part due to gas phase exothermic reactions



In the RPS region, N_2 molecules are mainly formed through surface recombination of N atoms with a probability

of 0.01, and so, the density of N_2 increases from the center of the RPS to the surface of the tube where recombination occurs, as well as increasing with the flow from the inlet to $4.0 \times 10^{12} \text{ cm}^{-3}$ at the end of the RPS. In the downstream chamber, the density of N_2 increases from 0.4 to $1.2 \times 10^{13} \text{ cm}^{-3}$ at the wafer surface as N_2 is formed in part through surface reactions in which abstraction of N(s) from the surface occurs [reactions in Eqs. (4a) and (6)]. The density of N_2 also increases through gas phase exothermic reactions of N with NF and NF_2 , mutual reactions of NF, and reactions of NO and N [reactions in Eqs. (9b), (10), (11b), and (11d)]. Once outside the region of either high T_e or high T_g , N_2 is fairly stable, tends to accumulate, and eventually flows out of the system.

Volatile SiF_4 is formed at the surface of the wafer through desorption of Si surface sites by successive passivation of F etchants in Eq. (3). The density of SiF_4 increases from $3.0 \times 10^{12} \text{ cm}^{-3}$ at the center of the wafer to $6.0 \times 10^{12} \text{ cm}^{-3}$ at the edge of the wafer as shown in Fig. 4(h) as the reaction products accumulate in the gas flow.

The densities of O_2 , O, FNO, and N_xO_y species (NO, NO_2 , and N_2O) are shown in Fig. 5. O_2 is dissociated in the RPS region through electron impact dissociative attachment and dissociative excitation



The density of O atoms peaks at $5.0 \times 10^{13} \text{ cm}^{-3}$ near the coils due to the locally high electron impact ionization source. In the downstream chamber, the density of O_2 increases from $2.0 \times 10^{13} \text{ cm}^{-3}$ near the showerheads to $3.0 \times 10^{13} \text{ cm}^{-3}$ at the wafer surface due to gas cooling. O_2 is also replenished through surface reactions abstracting O(s) sites [Eqs. (4c) and (4d)]. FNO is largely produced by the exothermic reaction of O and NF_2 [Eq. (9a)]. The increased density of FNO from $1.5 \times 10^{13} \text{ cm}^{-3}$ at the showerheads to $2.1 \times 10^{13} \text{ cm}^{-3}$ at the wafer surface as shown in Fig. 5(c) is due to decreased gas temperature which both favors the exothermic reaction between O and NF_2 [Eq. (9a)] over endothermic reactions and results in rebound in FNO density. The density of NO_2 has similar spatial variation with the flow as FNO, increasing from $3.0 \times 10^{11} \text{ cm}^{-3}$ at the showerheads to $1.0 \times 10^{12} \text{ cm}^{-3}$ at the wafer surface. NO_2 is predominantly formed through the exothermic reaction between FNO and O,



The formation of NO and N_2O is heavily modulated by the gas temperature and surface reactions. NO is mainly produced in the RPS region through endothermic reactions

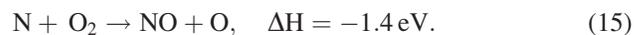


These reactions are aided by the high gas temperature (800–1000 K) in the RPS, and so the density of NO reaches its

maximum of $4.4 \times 10^{12} \text{ cm}^{-3}$ at the showerheads prior to the gas cooling. In the downstream chamber, NO is the main etchant to remove N surface sites by forming N_2 and N_2O through reactions in Eqs. (4a) and (5) and to remove O surface sites by forming NO_2 through the reaction in Eq. (4b). The density of NO monotonically decreases to $1.2 \times 10^{12} \text{ cm}^{-3}$ approaching the wafer, while the density of N_2O peaks at $2.5 \times 10^{12} \text{ cm}^{-3}$ at the wafer surface as it is an etch product. In addition to surface reactions, NO and N_2O are formed through gas phase exothermic reactions of N with NO_2 [Eqs. (11a) and (11c)] in the downstream chamber although the major source is by surface reactions.

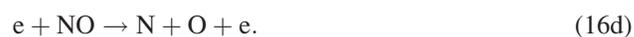
The fluxes to the wafer of the etchants of silicon nitride, F, NO, O, and N are shown in Fig. 6. As the ICP power increases from 100 to 800 W, the fractional dissociation of NF_3 increases from 34% to 88%, and the flux of F atoms to the center of the wafer increases from 0.6 to $1.7 \times 10^{18} \text{ cm}^{-2} \text{ s}^{-1}$. As the fractional dissociation of the NF_x species increases with power, the flux of F atoms to the wafer begins to saturate. There is only a nominal increase in the F flux when increasing the ICP power from 500 to 800 W. The flux of F atoms peaks at the center of the wafer due to the on-axis influx of reactants from the RPS.

As the power increases from 100 to 800 W, the gas temperature in the RPS increases from 630 to 1600 K, resulting in the enhanced formation of NO through endothermic reactions of N with O and O_2 [Eq. (14)] and the exothermic reaction having an activation energy of 3270 K,



The density of NO near the center of the wafer surface increases by an order of magnitude (from 0.2×10^{12} to $2.0 \times 10^{12} \text{ cm}^{-3}$), similar to the increase in the flux of NO to the wafer (0.3×10^{16} to $3.3 \times 10^{16} \text{ cm}^{-2} \text{ s}^{-1}$) while increasing ICP power from 100 to 800 W. The majority of this increase occurs between 100 and 300 W, when the gas temperature increases from 630 to 1080 K in the RPS. At these temperatures, endothermic reactions forming NO in the RPS begin to become effective. The flux of O atoms to the wafer surface almost linearly scales with the power, as only a single electron impact reaction with O_2 is required to produce O atoms by dissociative excitation and attachment [Eqs. (12a) and (12b)] while the initial O_2 is not significantly depleted.

The flux of N atoms to the wafer increases nonlinearly with increasing power, particularly above 300 W. N atoms are mainly formed through electron impact dissociative reactions of NF and NO,



As the ICP power increases, the production of NF through electron impact dissociative excitation of NF_2 [Eq. (7c)]

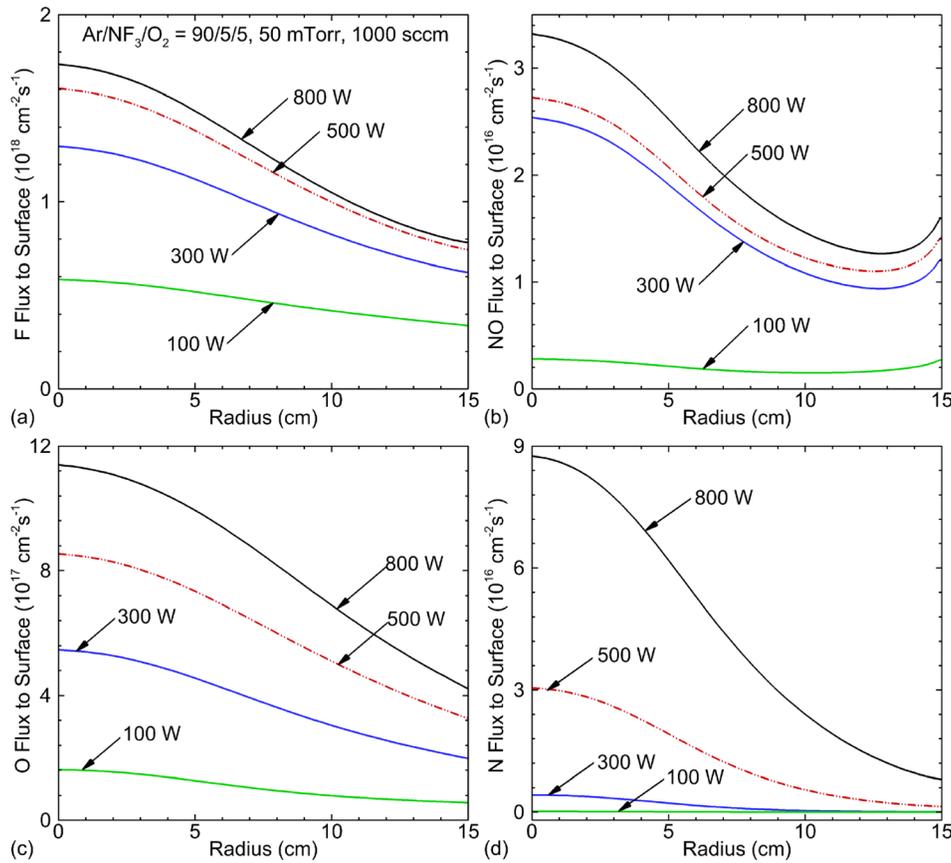


Fig. 6. (Color online) Time averaged fluxes of (a) F, (b) NO, (c) O, and (d) N to the surface of the silicon nitride wafer in a downstream etch system with a RPS driven by inductively coupled power. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, 10 MHz, and 100–800 W.

increases as NF₃ is more dissociated, while the production of NO through endothermic reactions increases due to the increase in gas temperature. The end result is that the flux of N atoms to the center of the wafer non-linearly increases from $1.8 \times 10^{14} \text{ cm}^{-2} \text{ s}^{-1}$ at 100 W to $8.7 \times 10^{16} \text{ cm}^{-2} \text{ s}^{-1}$ at 800 W. At 800 W, the flux of N atoms to the wafer is almost three times that of NO, resulting in a transition of the dominant reaction for the removal of N subsites from NO [Eqs. (4a) and (5)] to N [Eq. (6)].

The variations for the fractional coverage of Si(s) and N(s) subsites with ICP power are shown in Fig. 7, and the corresponding etch rates are shown in Fig. 8. Initially, the pristine silicon nitride surface consists of Si(s) subsites and N(s) subsites. The Si(s) subsites are rapidly removed by F atoms through the formation of volatile SiF₄ [Eq. (3)], leaving the N(s) subsites, which are then removed by NO and N radicals through reactions in Eqs. (4)–(6). For these conditions, the Si(s) subsites on the surface are almost completely etched due to the much larger flux of F atoms [etchant of Si(s) subsites] compared to fluxes of NO and N [etchants of N(s) subsites]. The complete etch process is rate limited by the availability of NO and N neutrals to remove N(s), leaving the top surface N(s) rich.

As the ICP power increases from 100 to 800 W, the non-linear increase of fluxes of NO and N shown in Fig. 6 favors the more rapid removal of N(s) subsites compared to the removal of Si(s) sites. These trends result in an increase in

the etch rate (from 20 to 320 nm/min at the wafer center and from 20 to 90 nm/min at the wafer edge) shown in Fig. 8. The exposure of Si(s) subsites increases, while the fractional coverage of N(s) subsites at the center of the wafer decreases from 0.95 to 0.74, as shown in Fig. 7(b). The fractional coverage of O(s) sites is less than 0.02 as the O flux which etches those sites is larger than the NO flux by an order of magnitude. Once O(s) surface sites are formed through surface oxidation by NO molecules in Eq. (4a), they are quickly etched by gas phase O atoms through the formation of O₂ in Eq. (4c).

The ratio of fluxes of F atoms to NO molecules and N atoms, $\phi_F/(\phi_{\text{NO}} + \phi_{\text{N}})$, is a key parameter for determining the etch rate and uniformity. For similar reaction probabilities, the ratio of the etchants of Si(s) subsites to the etchants of N(s) subsites should align with the stoichiometry of silicon nitride. The limiting factor for the etch rate in the F rich condition is then the NO and N fluxes which remove N(s) surface sites. Thus, the etch rate nearly linearly increases with the fluxes of NO and N radicals to the wafer surface as shown in Fig. 8(b). The center peaked etch rates at the higher ICP powers can be attributed to the nonuniform fluxes of N and NO reaching the surface, which is due to radial diffusion of N and NO in the downstream chamber as shown in Figs. 4(f) and 5(e).

In the experiments of downstream etching of silicon nitride by a RPS sustained in NF₃/O₂ performed by

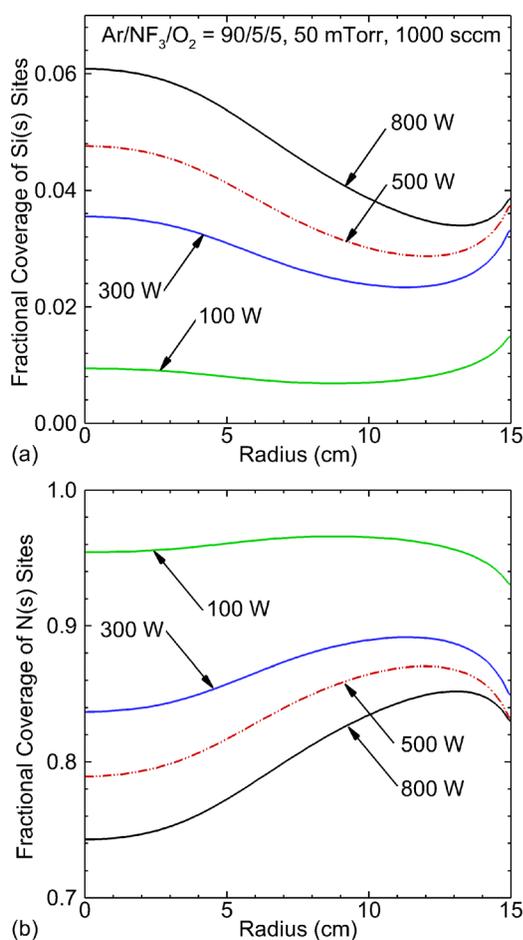


FIG. 7. (Color online) Fractional coverage of (a) Si and (b) N subsites at the surface of the silicon nitride wafer in a downstream etch system with a RPS driven by inductively coupled power. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, 10 MHz, and 100–800 W.

Kastenmeier *et al.*, the etch rate and degree of surface oxidation were proportional to the partial pressure of NO in the reactor chamber.^{4,22} The small contribution of N atoms to the etch rate in their experiments was due to the small flux of N atoms to the wafer for their operating conditions.

With the goal of this investigation being elucidating reaction mechanisms, an exhaustive study of methods to improve etch uniformity was not performed. The primary source of nonuniformities is the on-axis nozzle and the relatively short downstream diffusion chamber. It is more typical in industrial reactors to use a plenum, a wafer-diameter showerhead, and a taller downstream chamber to disperse the reactants more uniformly over the wafer. As a demonstration, limited studies were performed using a plenum connecting the RPS and the downstream reactor using the geometry shown in Fig. 9. The height of the plenum is 4 cm, and the radius is 13 cm. The flow distance from the RPS to the wafer is 17 cm with two levels of showerheads placed in-between—one at the end of the confined plasma zone and the other at the base of the plenum. The spacing of the showerheads (0.75 cm), the ICP power (300 W), and the other operating conditions are the same as the base case.

Gas temperature and reactant densities for the RPS with plenum are shown in Fig. 9. There is a significant gradient in

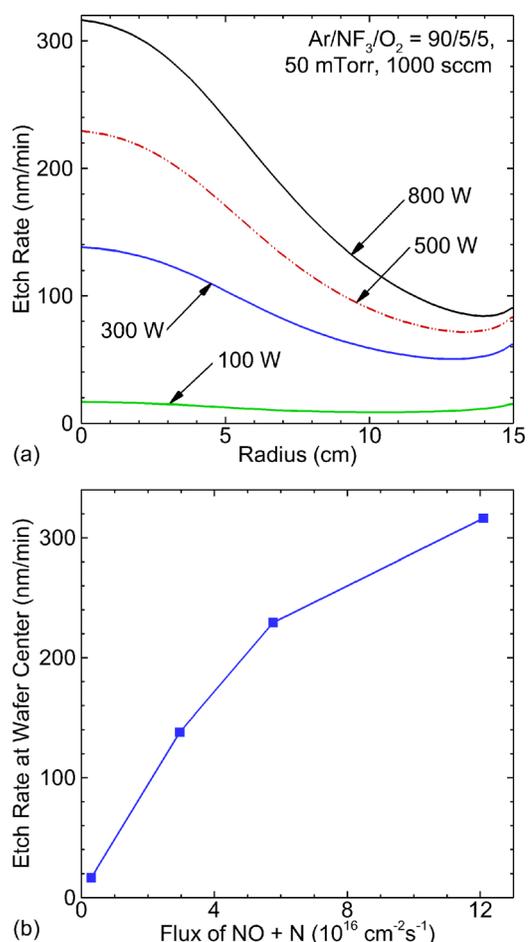


FIG. 8. (Color online) Time averaged etch rates of silicon nitride (a) along the radius of the wafer at different ICP powers and (b) at the center of the wafer surface for different fluxes of NO and N radicals in a downstream etch system with a RPS driven by inductively coupled power. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, 10 MHz, and 100–800 W.

the gas temperature from the source region (≈ 1100 K) to the plenum (≈ 380 K) and the downstream chamber (≈ 330 K). This gradient provides opportunities to customize the gas reaction pathways and optimize radical production through combinations of endothermic and exothermic reactions by properly choosing the injection points for process gases. In the plenum, the neutrals diffuse more in the radial direction than in the vertical direction due to the higher pressure and resistance to the gas flow produced by the showerheads placed between the plenum and the downstream chamber. The densities of F and O radicals near the circumference of the plenum are about 60% of the values at the axis of the plenum. The density of N atoms decreases from $1.0 \times 10^{12} \text{ cm}^{-3}$ at the plenum axis to $0.1 \times 10^{12} \text{ cm}^{-3}$ near the fringe, while the density of NO only slightly decreases from 5.8×10^{12} to $5.4 \times 10^{12} \text{ cm}^{-3}$. These trends are due to exothermic reactions of N with NO₂ and O₂ [Eqs. (11a) and (15)], which consume N atoms and produce NO molecules.

The neutral flow from the plenum into the downstream chamber through the array of showerheads produces more uniform densities in the downstream chamber, shown in Fig. 9, and more uniform fluxes to the wafer, shown in Fig. 10.

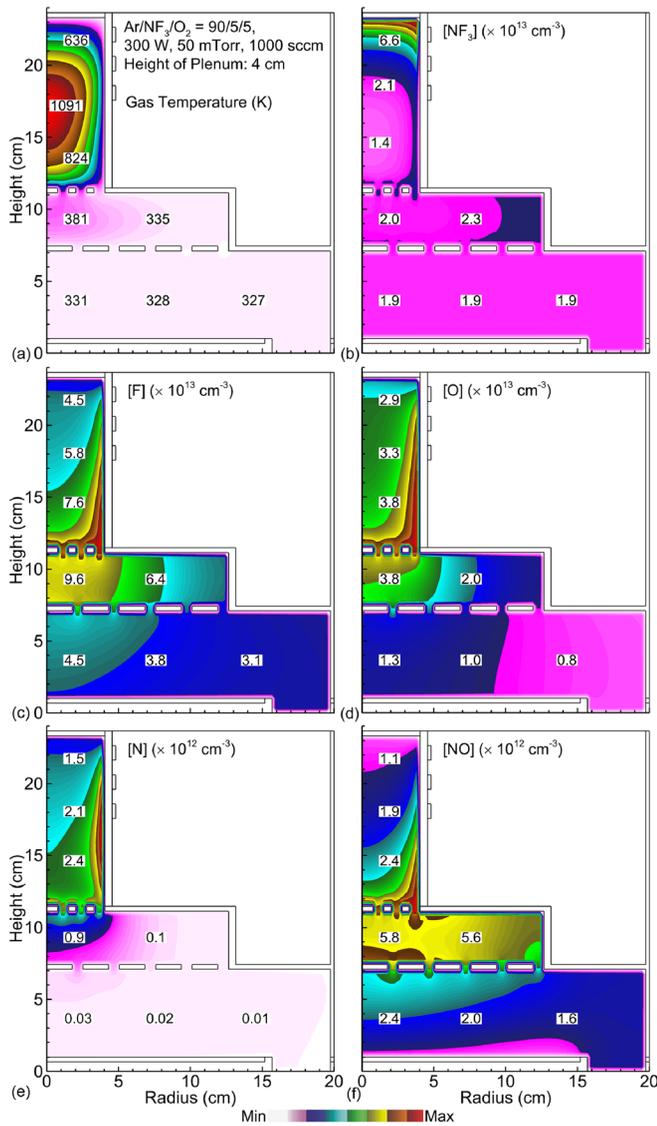


FIG. 9. (Color online) Time averaged (a) gas temperature and densities of (b) NF_3 , (c) F, (d) O, (e) N, and (f) NO in a downstream etch system consisting of a RPS driven by inductively coupled power, a plenum with a height of 4 cm, and a downstream reactor. Operating conditions: $\text{Ar}/\text{NF}_3/\text{O}_2 = 90/5/5$, 50 mTorr, 1000 sccm, 300 W, and 10 MHz.

With the gas temperature in the downstream chamber being nearly uniformly 330 K, there is no thermal radial gradient across the wafer, which could otherwise produce nonuniform neutral fluxes. The flux of F decreases from $6.2 \times 10^{17} \text{ cm}^{-2} \text{ s}^{-1}$ at the center of the wafer to $4.3 \times 10^{17} \text{ cm}^{-2} \text{ s}^{-1}$ at the edge of the wafer, and the flux of NO decreases from 1.4×10^{16} to $1.0 \times 10^{16} \text{ cm}^{-2} \text{ s}^{-1}$, about a 30% difference compared to about 60% without the plenum.

As the flux of F to the wafer is 1 order of magnitude larger than that of NO and N, the Si(s) subsites are almost totally removed, resulting in exposure of N(s) subsites, and so, the uniformity of etching is mainly determined by the uniformity of the flux of NO and N to the surface. The etch rate shown in Fig. 10(c) decreases from 70 nm/min at the wafer center to 50 nm/min at the edge (30% difference), providing better uniformity than the base case without the

plenum [Fig. 8(a)]. Although the addition of the plenum improves the uniformity of etching, the overall etch rate decreases due to the longer flow time (stagnation in the plenum), which affords more opportunity for recombination reactions of radicals and more opportunity for surface reactions with chamber walls. Uniformity can be further improved by increasing the volume of the plenum (at the cost of reducing the radical density), having a hole-pattern in the showerhead with larger diameter holes at a larger radius

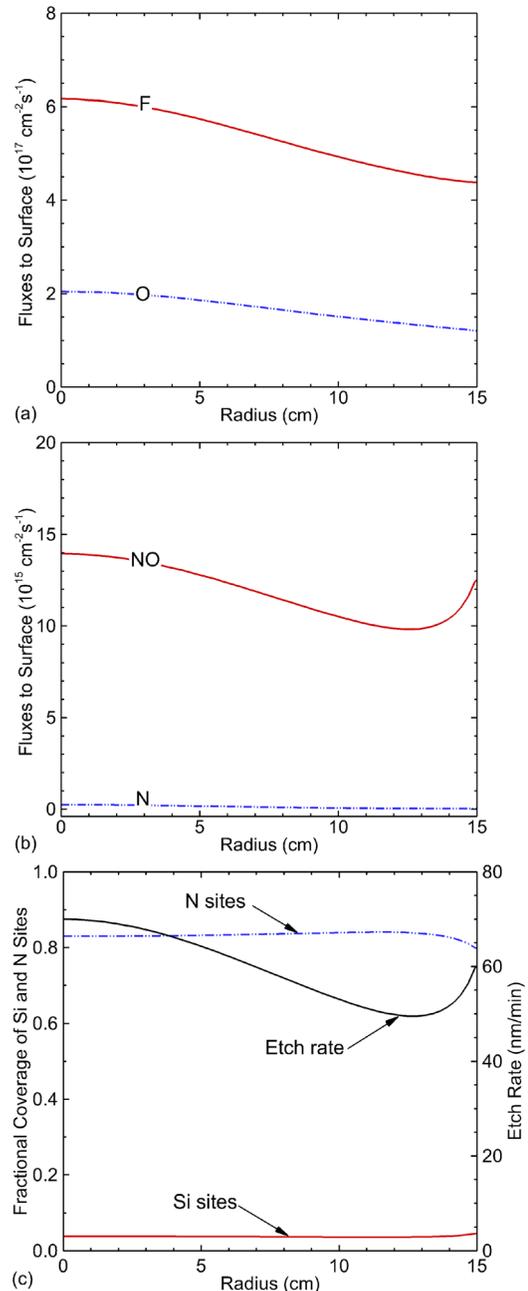


FIG. 10. (Color online) Time averaged fluxes of (a) F and O and (b) NO and N; (c) fractional coverage of Si and N subsites and etch rate of the silicon nitride wafer in a downstream etch system consisting of a RPS driven by inductively coupled power, a plenum with a height of 4 cm, and a downstream reactor. Operating conditions: $\text{Ar}/\text{NF}_3/\text{O}_2 = 90/5/5$, 50 mTorr, 1000 sccm, 300 W, and 10 MHz.

and by increasing the distance between the showerhead and the wafer (again, at the cost of reducing radical density).

V. REMOTE PLASMA SOURCE SUSTAINED BY PULSED INDUCTIVELY COUPLED POWER

Pulsing enables both plasma and etch properties to be controlled beyond what may be possible with continuous wave (CW) excitation. For CW excitation, the electron source and loss need to be instantaneously balanced, producing a unique T_e for a given pressure and power. As a result, the balance between generating radicals and ions is fairly constrained. The use of pulsed power provides a means to customize the electron energy distribution by allowing the electron source and loss to be balanced over the pulse period as opposed to instantaneously.²⁸ The electron temperature can considerably vary during the pulse period. Although the gas temperature can also vary during a pulse period due to modulation of the gas heating sources by electron impact dissociative reactions, the degree of modulation is less due to the finite heat capacity of the system. The result is some ability to customize rate coefficients for both electron impact reactions and heavy particle endothermic and exothermic reactions. For systems with the wafer in direct contact with the plasma, the ion-ion plasma formed during the afterglow minimizes damage to the surface by reducing net charge accumulation on the wafer.²⁹

The RPS was sustained with pulsed 10 MHz inductively coupled power with a 25 kHz pulse repetition frequency (PRF) and 25% DC. (DC refers to the fraction of the pulse period that power is applied, and PRF refers to the number of times per second the waveform is repeated.) The pulsed-period-averaged (PPA) power was 300 W. The other operating conditions were kept the same as those for the base CW case (50 mTorr, Ar/NF₃/O₂ = 90/5/5, 1000 sccm). The variations of volume averaged densities of charged particles and electron temperature with time are shown in Fig. 11. (These densities are volume averaged weighted by the electron density and so preferentially show densities in the RPS.) These values are shown after 26 pulsed periods so that a pulse-periodic steady state is achieved.

With a fast rising power pulse, increasing power from 0 to 1200 W in 1.5 μ s delivered into an initially small inventory of electrons having a density of $\approx 1.0 \times 10^9 \text{ cm}^{-3}$ produces a spike of electron temperature to 5.4 eV. This high T_e above the CW value (4.2 eV) and the higher power during the pulse enable more efficient excitation and ionization, which leads to an increased electron density of $5.0 \times 10^{10} \text{ cm}^{-3}$ compared to CW. During the pulse-on period, charge neutrality is mainly maintained by $[e] + [F^-] \approx [Ar^+] + [O_2^+]$. The dominant positive ions are Ar⁺ and O₂⁺ due to their large initial mole fractions. Due to low mole fractions (<5%) and moderate ionization potentials (11–13 eV) of NF_x, the densities of NF_x⁺ ($\sim 10^8 \text{ cm}^{-3}$) are 2 orders of magnitude lower than Ar⁺. When the power is turned off, the electrons quickly attach to NF_x to form F⁻ or diffuse to the walls, resulting in the transition of the dominant negative charge from electrons to F⁻ ions. The loss mechanisms for the positive ions in the

afterglow are recombination with electrons (20% of the total loss), neutralization with negative ions (80%), and diffusion to the walls (<1%). As the ionization potentials of O₂ (12.1 eV) and NO (9.3 eV) are lower than those of Ar (16.0 eV) and O (16.2 eV), O₂⁺ and NO⁺ are replenished by charge exchange and Penning ionization in the afterglow and their densities have a slower decay rate compared to other positive charged species.

The spatial distributions of electron density, electron temperature, and electron impact ionization source in the middle of the pulse-on period and the middle of the pulse-off period are shown in Fig. 12. During the pulse-on period, the electron density and electron temperature peak at $2.6 \times 10^{12} \text{ cm}^{-3}$ and 4.5 eV adjacent to the coils in the electromagnetic skin depth. The electron impact ionization source peaks at $3.6 \times 10^{17} \text{ cm}^{-3} \text{ s}^{-1}$ near the coils, which is about six times that of the CW value shown in Fig. 2(c). During the pulse-off period, the electron density and electron temperature drop to 10^{10} cm^{-3} and below 0.3 eV. At these temperatures, thermal attachment is favored over the ionization, resulting in S_e transitioning from positive in Fig. 12(a) (with regions of the negative source away from the coil) to

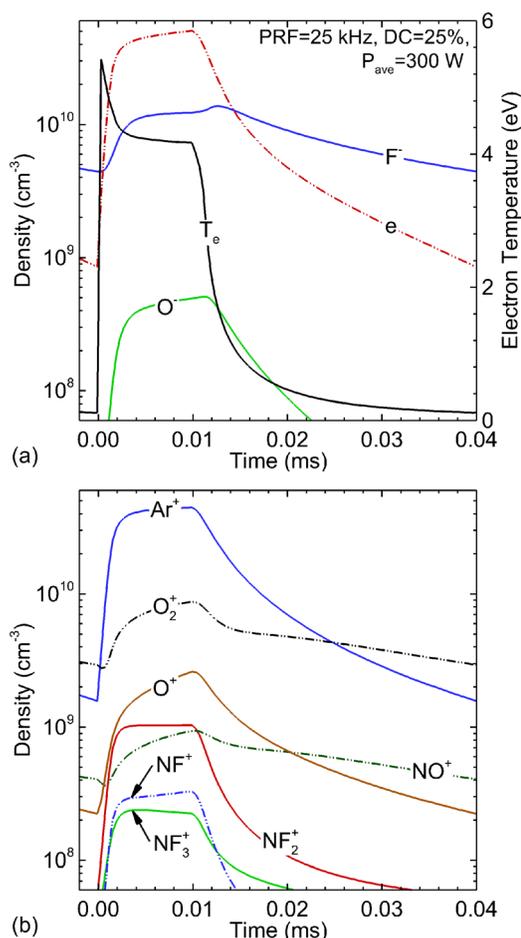


Fig. 11. (Color online) Time variations for the volume averaged (a) electron density, electron temperature, and densities of negative ions and (b) densities of positive ions in a downstream etch system with a RPS driven by pulsed inductively coupled power. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, PPA power: 300 W, PRF = 25 kHz, and DC = 25%.

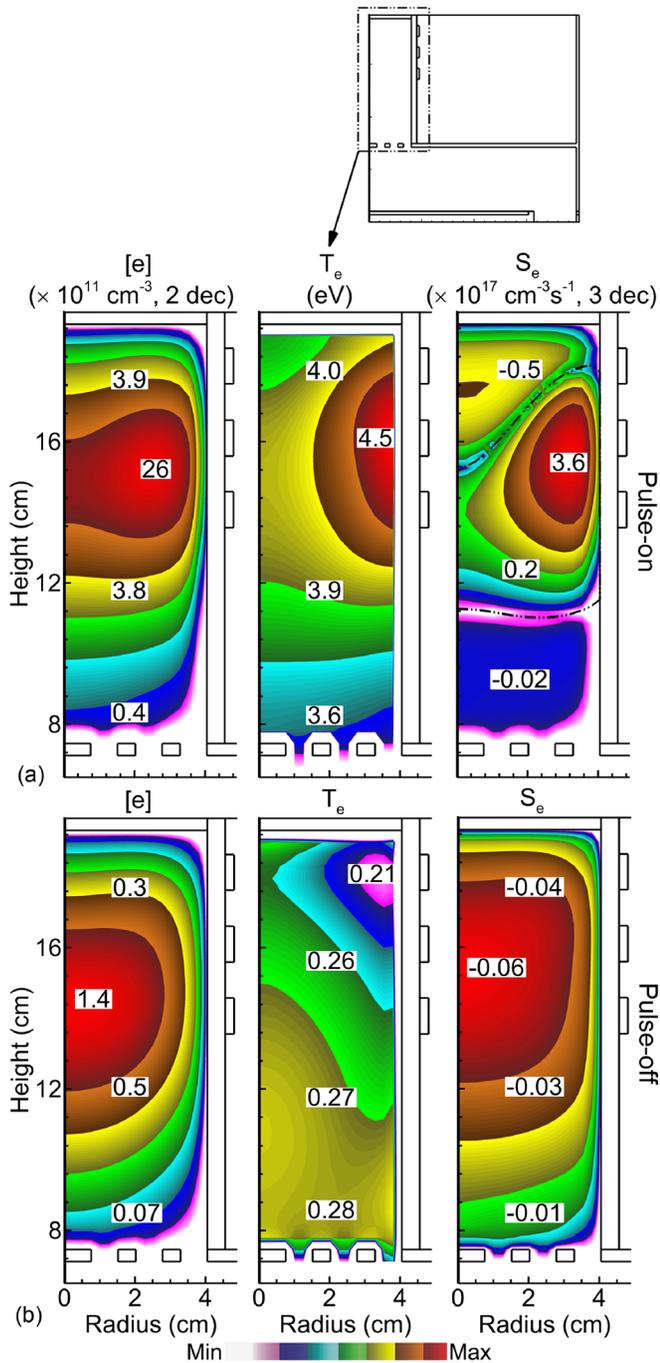


Fig. 12. (Color online) Electron density, electron temperature, and electron impact ionization source in the middle of (a) pulse-on and (b) pulse-off periods in a RPS driven by pulsed inductively coupled power. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, PPA power: 300 W, PRF = 25 kHz, and DC = 25%.

fully negative in Fig. 12(b). The maximum net loss of $6.2 \times 10^{15} \text{ cm}^{-3} \text{ s}^{-1}$ occurs at the center of the RPS reactor where the electron density is the largest in the afterglow.

The densities of F⁻ and NO⁺ are shown in Fig. 13 during the power-on and power-off portions of the cycle. These ions are only moderately modulated by the pulsed power in the RPS region, being a factor of 2 larger during the pulse, whereas the electron density is modulated by an order of magnitude. The time for ion-ion recombination and for

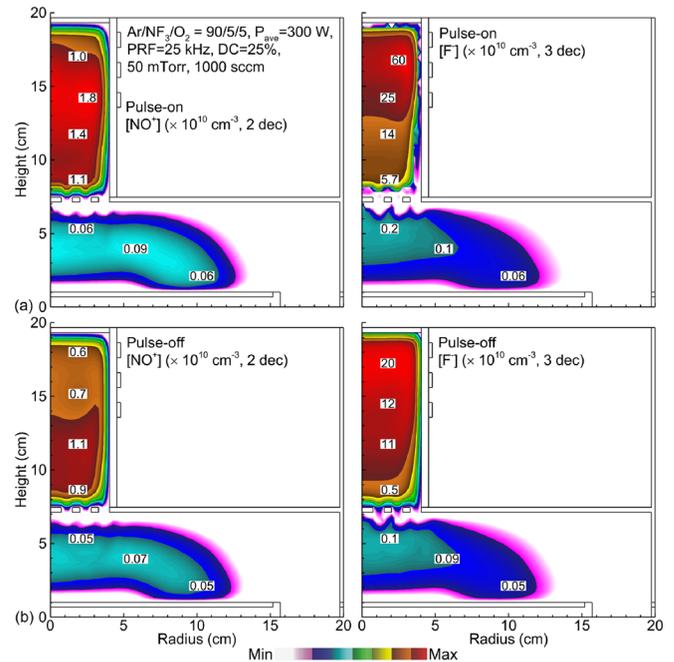


Fig. 13. (Color online) Densities of NO⁺ and F⁻ in the middle of (a) pulse-on and (b) pulse-off periods in a downstream etch system with a RPS driven by pulsed inductively coupled power. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, PPA power: 300 W, PRF = 25 kHz, and DC = 25%.

diffusion to the walls is longer than the interpulse period. With the low electron temperature and low electron density, there is little ambipolar enhancement to diffusion rates. The densities of F⁻ and NO⁺ ions are also only slightly modulated in the downstream chamber, where charge neutrality is maintained by $[\text{NO}^+] \approx [\text{F}^-]$. The showerheads serve well to isolate the ions that have leaked into the downstream chamber from the dynamics occurring in the RPS. The ambipolar electric fields in the downstream chamber are small since the ion charge density is only marginally at the level of supporting ambipolar electric fields. As a result, the ions nearly freely diffuse. What modulation does occur depends on greater injection (or leakage) of ions from the RPS during the pulse-on period. The electron density in the downstream chamber is below 10^7 cm^{-3} , whereas the ion-ion plasma has a density of 10^8 cm^{-3} . As in the CW case, there is negligible flux of ions to the wafer surface and the etching of the silicon nitride wafer is almost purely neutrally driven.

To investigate the effect of different strategies of pulsed power on the plasma and etch properties, the PRF was varied from 25 to 100 kHz and the DC was varied from 25% to 75%, while the PPA power was 300 W. The volume averaged densities of electrons and ions and electron temperatures for different DCs and PRFs are shown in Figs. 11, 14, and 15. As the DC decreases from 75% to 25% with the PRF maintained at 25 kHz, the peak power increases from 400 to 1200 W. Over this range, the electron density in the pulse-on period increases from 1.8×10^{10} to $5.0 \times 10^{10} \text{ cm}^{-3}$ and the peak electron temperature increases from 5.2 to 5.4 eV, resulting in the enhanced electron impact ionization source and increased densities of positive ions by 2–3 times as

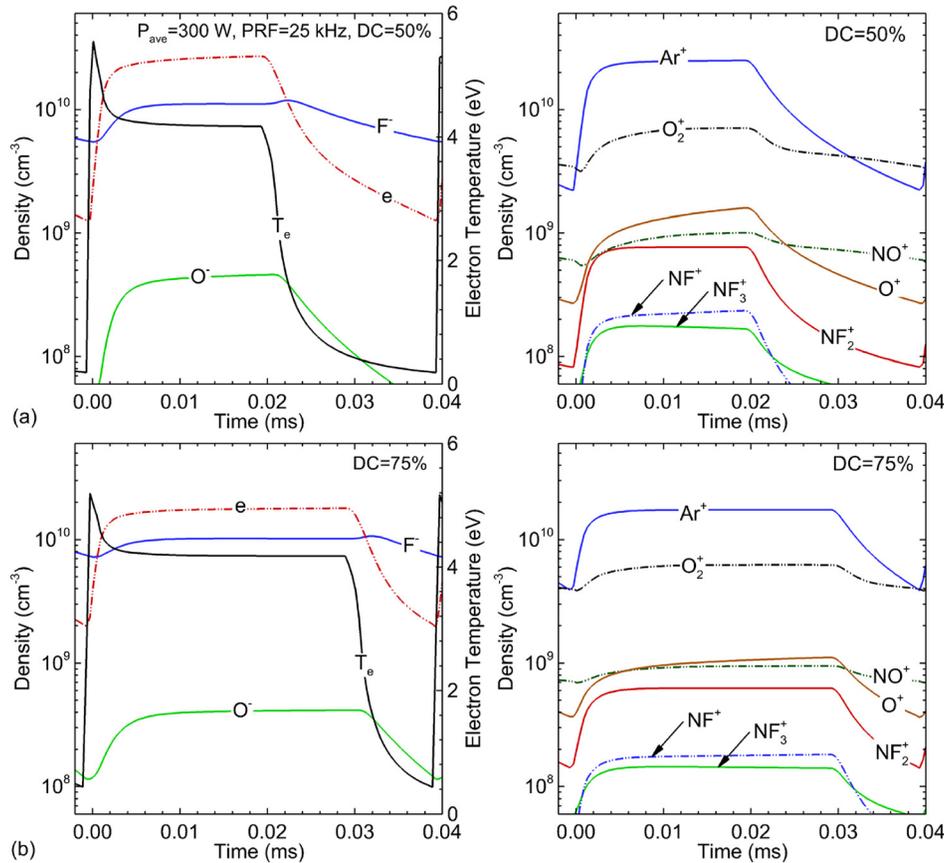


Fig. 14. (Color online) Time variations for the volume averaged electron density, electron temperature, and ion densities in a downstream etch system with a RPS driven by pulsed inductively coupled power with duty cycles of (a) 50% and (b) 75%. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, PPA power: 300 W, PRF = 25 kHz, and DC = 50%, 75%.

shown in Figs. 11 and 14. The electron loss through diffusion, attachment, and recombination becomes larger with decreasing DC due to the longer pulse-off period, and so, the electron density at the end of the pulse-off period decreases from $2 \times 10^9 \text{ cm}^{-3}$ for DC = 75% to $1 \times 10^9 \text{ cm}^{-3}$ for DC = 25%. With this reduced conductivity at the end of the pulse-off period, there is proportionally more collisional heating at the beginning of the pulse-on period, with a larger E/N in the bulk plasma.

As the PRF increases from 25 to 100 kHz with the DC maintained at 25%, the electron loss during the pulse-off period becomes smaller due to the shorter pulse-off period, leading to higher conductivity and lower E/N at the beginning of the next pulse-on period. As a result, the peak electron temperature decreases from 5.4 to 5.2 eV as shown in Figs. 11 and 15. As the PRF varies, the plasma properties during the pulse-on period are slightly modulated as the peak power is the same (1200 W) due to the same DC (25%) and PPA power (300 W).

As radical lifetimes and the residence time of the flow (\sim ms) are much longer than the pulse period (10–40 μ s), varying the PRF (with constant DC) from 25 to 100 kHz results in little modulation in the densities and fluxes of neutrals to the wafer, and so, the etch properties (e.g., etch rate and uniformity) are almost invariant when varying the PRF.

The pulse averaged fluxes of the etchants for silicon nitride (F, NO, O, and N) to the wafer surface for different DCs for a constant PRF (25 kHz) are shown in Fig. 16. The pulse averaged fractional coverage of Si(s) and N(s) subsites and etch rates are shown in Fig. 17. The pulse averaged flux of F atoms only slightly increases with decreasing DC from 100% (base CW case) to 25% as shown in Fig. 16(a). This small increase is due to the enhanced dissociation by the spike in electron temperature (5.2–5.4 eV) during pulsed cases above the CW value (4.2 eV). However, since the majority of F atoms are produced through dissociative attachment to NF_x by thermal electrons, this spiking of the electron temperature does not significantly affect the F atom production.

The pulse averaged flux of NO to the center of the wafer decreases from 2.5×10^{16} to $1.8 \times 10^{16} \text{ cm}^{-2} \text{ s}^{-1}$ with decreasing DC from 100% to 25% as shown in Fig. 16(b). NO is primarily formed through the endothermic reactions of N₂ with O and O₂ [Eq. (14)] and exothermic reactions of N with O₂ [Eq. (15)], with the latter having an activation energy of 3270 K. The density of NO is therefore highly dependent on excursions of the gas temperature which approach the activation energy. As the DC decreases from 100% to 25% and peak power increases, the volume averaged gas temperature in the RPS during the pulse-on period

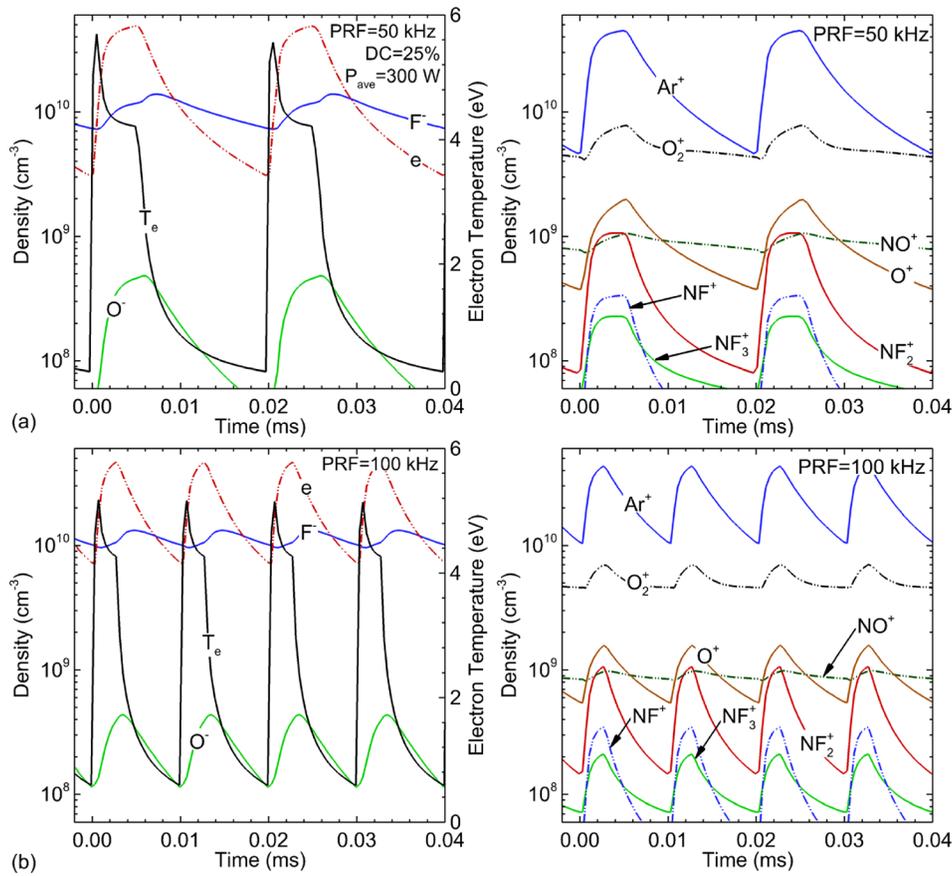


FIG. 15. (Color Online) Time variations for the volume averaged electron density, electron temperature, and ion densities in a downstream etch system with a RPS driven by pulsed inductively coupled power with PRF of (a) 50 kHz and (b) 100 kHz. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, PPA power: 300 W, PRF = 50, 100 kHz, and DC = 25%.

increases from 850 to 875 K, while the peak temperature increases from 1080 to 1150 K. However, during the pulse-off period, T_g decreases from 850 to 635 K due to the longer period for thermal conduction to the walls. Overall, the enhancement of endothermic reactions for the formation of NO during the pulse-on period cannot compensate for losses due to exothermic reactions in the pulse-off period. As a result, the density of NO near the wafer surface decreases from 1.2×10^{12} to $0.9 \times 10^{12} \text{ cm}^{-3}$ with decreasing DC, and the flux of NO decreases as shown in Fig. 16(b).

The flux of F atoms to the surface monotonically decreases from the center to the edge of the wafer, while a minimum occurs for the flux of NO near the edge of the wafer at a radius of 13 cm shown in Fig. 16(b). NO molecules are consumed by N(s) subsites to form N₂O. As the N₂O etch product diffuses from the wafer, the small increase in the flux of NO at the edge of the wafer is then due to the exothermic reaction



The electron density and electron temperature during the pulse-on period increase as DC decreases, which favors the formation of O atoms by electron impact dissociation of O₂ [Eq. (12)] and the formation of N atoms through electron impact dissociation of NF and NO [Eq. (16)]. Nevertheless,

the pulse averaged flux of O to the center of the wafer surface decreases from 5.4×10^{17} to $4.8 \times 10^{17} \text{ cm}^{-2} \text{ s}^{-1}$ and that of N decreases from 4.2×10^{15} to $1.6 \times 10^{15} \text{ cm}^{-2} \text{ s}^{-1}$. This trend is due to the increased production at lower DC not being able to compensate for the consumption of N and O during the longer afterglow.

The fractional coverages of Si(s) and N(s) subsites and etch rate as a function of DC are shown in Fig. 17. In fluorine rich conditions where the density of F atoms is 2 orders of magnitude larger than the densities of NO and N, the Si(s) subsites are almost completely removed by the F atoms through reaction in Eq. (3), resulting in a small (<0.04) coverage of Si(s) subsites. As the fluxes of NO and N to the surface increase with increasing DC, the removal of N(s) subsites and re-exposure of underlying Si(s) subsites are favored. This leads to a small decrease in fractional coverage of N(s) subsites at the center of the surface from 0.88 to 0.84 and increased fractional coverage of Si(s) subsites from 0.025 to 0.035. The etch rate at the center of the surface increases from 96 to 138 nm/min with increasing DC. For fluorine rich conditions, the etch rate is limited by the availability of the etchants of N(s) subsites (i.e., NO and N). When increasing DC, the densities and fluxes of both NO and N increase, which being the rate limiting flux for completing the etch cycle increases the etch rate.

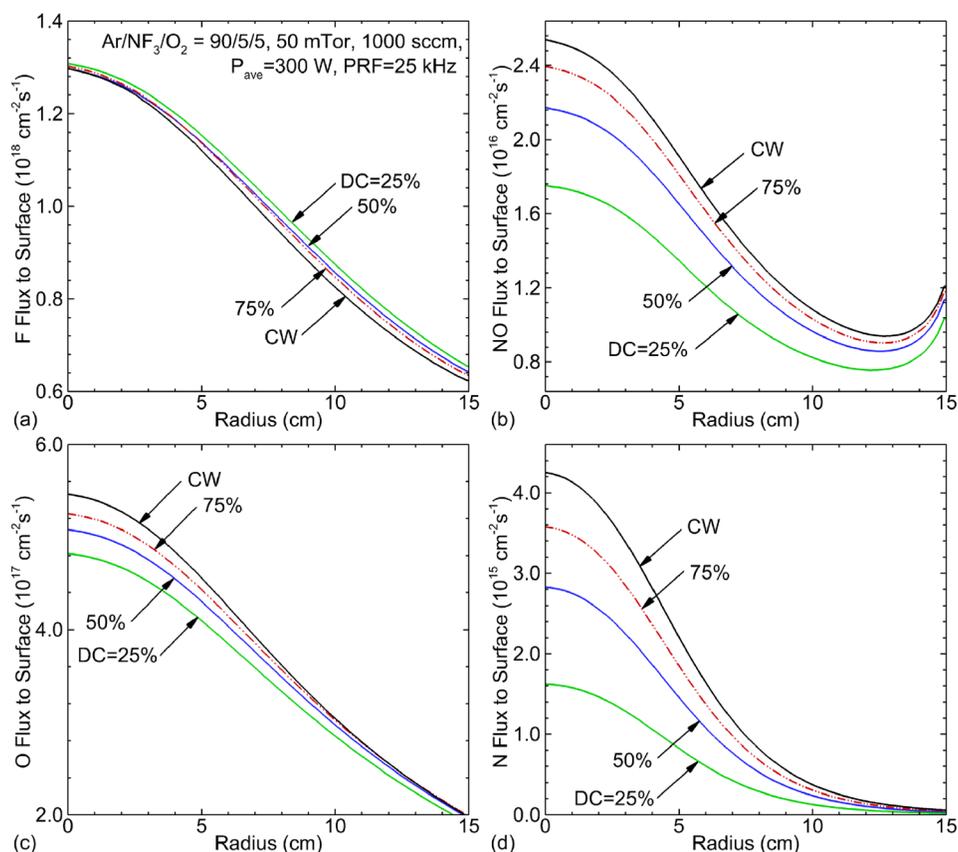


FIG. 16. (Color online) Pulse averaged fluxes of (a) F, (b) NO, (c) O, and (d) N to the surface of the silicon nitride wafer in a downstream etch system with a RPS driven by pulsed inductively coupled power. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, PPA power: 300 W, PRF = 25 kHz, and DC = 25%–100%.

VI. CONCLUDING REMARKS

A downstream plasma etching system consisting of a RPS sustained in Ar/NF₃/O₂ mixtures with CW and pulsed ICP power and a downstream chamber with a silicon nitride coated wafer placed on the substrate has been modeled using a two-dimensional simulation. In the RPS, the dominant positive ions are Ar⁺ and O₂⁺ due to the large mole fraction of their parent gases and the dominant negative ions are F⁻ due to thermal attaching NF_x species. Most of the ions are confined in the source region, except that a weak ion-ion plasma maintained by [NO⁺] ≈ [F⁻] persists into the downstream chamber due to the lowest ionization potential being that of NO and the highest electron affinity being that of F among all the important neutral species in the system. As negligible fluxes of ions reach the surface of the wafer, etching is almost purely neutrally driven and so is nearly damage free and isotropic.

A surface reaction mechanism for the etching of silicon nitride by thermal neutrals was developed and implemented into the two-dimensional model. The etching proceeds by iteratively removing Si and N surface subsites. Si subsites are removed through the formation of volatile SiF₄ by successive passivation by F atoms, while N subsites are etched through the formation of N₂O and N₂ by NO and N. As the ICP power increases for CW excitation, more F and N atoms are produced through electron impact dissociative reactions

of NF_x and NO, leading to increased gas temperature due to Franck–Condon heating, which in turn favors the formation of NO through endothermic reactions in the source region. In fluorine rich conditions, the etch rate is limited by the availability of the etchants of N subsites, NO and N, and so, the etch rate nearly linearly increases with fluxes of NO and N to the wafer.

The downstream etching process has been investigated and mildly optimized, using two approaches. From the perspective of geometry, a plenum added between the remote plasma source and the downstream chamber favors transverse diffusion of neutrals and results in neutral fluxes to the wafer and etch rates of silicon nitride having higher uniformity. From the perspective of excitation power, the use of pulsed power source provides a means to customize the electron impact reactions and heavy particle endothermic and exothermic reactions by enabling different electron temperatures and gas temperatures over the pulse period. As the powered portion of the cycle increases with the increasing duty cycle, pulse averaged fluxes of NO and N to the wafer increase, resulting in more removal of N subsites and thus increased etch rates for these fluorine rich conditions.

In addition to using a plenum and pulsed power sources, other optimization strategies are suggested by these results. The etch rate can be better controlled by separately controlling the production of F and N_xO_y species using multiple

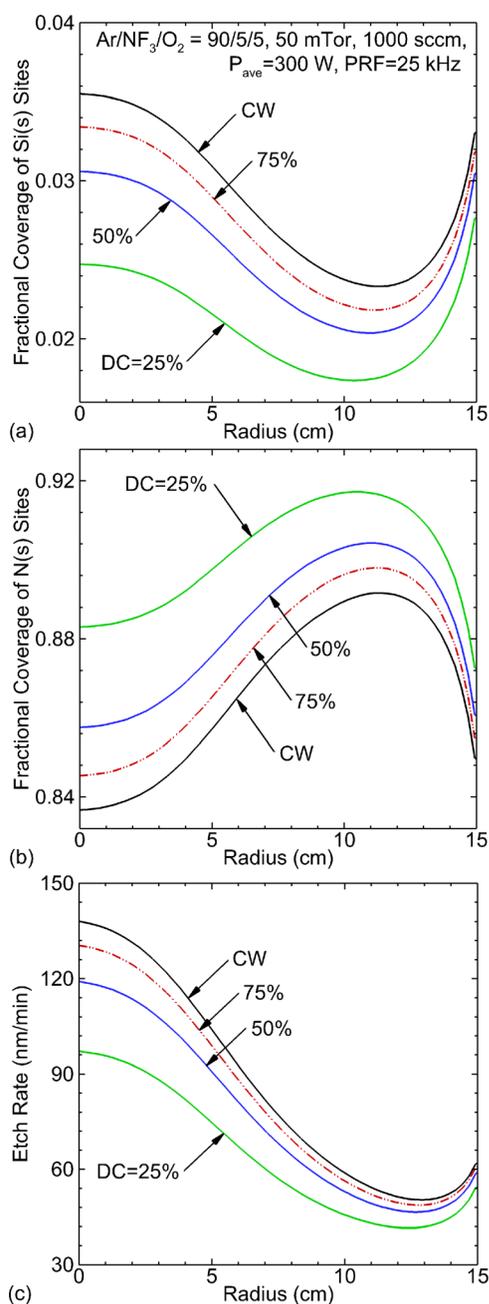


Fig. 17. (Color Online) Pulse averaged fractional coverage of (a) Si and (b) N subsites at the wafer surface and (c) etch rates of silicon nitride in a downstream etch system with a RPS driven by pulsed inductively coupled power. Operating conditions: Ar/NF₃/O₂ = 90/5/5, 50 mTorr, 1000 sccm, PPA power: 300 W, PRF = 25 kHz, and DC = 25%–100%.

remote sources. In this way, the ratio of the etchants of Si subsites (F atoms) to the etchants of N subsites (NO and N) reaching the wafer aligns with the stoichiometry of silicon nitride. In downstream etching systems, the gas temperature can exceed several thousand degrees in the RPS while decreasing to near ambient in the downstream chamber. This natural gradient in gas temperature may enable customizing the reaction pathway using multiple gas inlets at different gas temperatures (for example, one inlet at the upstream of

RPS and the other at the downstream of RPS). The endothermic reactions mainly occur in the RPS, uniquely producing precursors which flow downstream and exothermically react with process gases injected from the downstream inlet to produce reactants for etching and surface modification at the wafer surface.

ACKNOWLEDGMENTS

This work was supported by the Samsung Electronics Co. Ltd., the DOE Office of Fusion Energy Science (Nos. DE-SC0001319 and DE-SC0014132), and the National Science Foundation (Nos. PHY-1519117 and PHY-1500126).

- ¹N. Posseme, V. Ah-Leung, O. Pollet, C. Arvet, and M. Garcia-Barros, *J. Vac. Sci. Technol., A* **34**, 061301 (2016).
- ²Q. Qian, Z. Zhang, M. Hua, G. Tang, J. Lei, F. Lan, Y. Xu, R. Yan, and K. J. Chen, *Nanotechnology* **28**, 175202 (2017).
- ³V. M. Donnelly and A. Kornblit, *J. Vac. Sci. Technol., A* **31**, 050825 (2013).
- ⁴B. E. E. Kastenmeier, P. J. Matsuo, G. S. Oehrlein, and J. G. Langan, *J. Vac. Sci. Technol., A* **16**, 2047 (1998).
- ⁵B. E. E. Kastenmeier, P. J. Matsuo, J. J. Beulens, and G. S. Oehrlein, *J. Vac. Sci. Technol., A* **14**, 2802 (1996).
- ⁶X. Liu, S. Gill, F. Tang, S. W. King, and R. J. Nemanich, *J. Vac. Sci. Technol., B* **30**, 031212 (2012).
- ⁷S. C. Heo, D. Lim, W. S. Jung, R. Choi, H.-Y. Yu, and C. Choi, *Microelectron. Eng.* **147**, 239 (2015).
- ⁸H. Tanaka *et al.*, *IEEE Symposium on VLSI Technology* (2007), p. 14.
- ⁹J. Jang *et al.*, *IEEE Symposium on VLSI Technology* (2009), p. 192.
- ¹⁰M. G. Blain, T. L. Meisenheimer, and J. E. Stevens, *J. Vac. Sci. Technol., A* **14**, 2151 (1996).
- ¹¹T. Hayashi, K. Ishikawa, M. Sekine, M. Hori, A. Kono, and K. Suu, *Jpn. J. Appl. Phys.* **51**, 016201 (2012).
- ¹²J. Kikuchi, M. Iga, H. Ogawa, S. Fujimura, and H. Yano, *Jpn. J. Appl. Phys., Part 1* **33**, 2207 (1994).
- ¹³J. Lee, S. J. Lee, W. B. Han, H. Jeon, J. Park, W. Jang, C. S. Yoon, and H. Jeon, *Phys. Status Solidi A* **210**, 276 (2013).
- ¹⁴H. Zhu, X. Qin, L. Cheng, A. Azcatl, J. Kim, and R. M. Wallace, *ACS Appl. Mater. Interfaces* **8**, 19119 (2016).
- ¹⁵T. Hayashi, K. Ishikawa, M. Sekine, M. Hori, A. Kono, and K. Suu, *Jpn. J. Appl. Phys.* **51**, 026505 (2012).
- ¹⁶H. Nishino, N. Hayasaka, and H. Okano, *J. Appl. Phys.* **74**, 1345 (1993).
- ¹⁷E. Meeks, R. S. Larson, S. R. Vosen, and J. W. Shon, *J. Electrochem. Soc.* **144**, 357 (1997).
- ¹⁸S. R. Vosen, E. Meeks, R. S. Larson, and J. W. Shon, *J. Electrochem. Soc.* **144**, 1514 (1997).
- ¹⁹Y. Barsukov, V. Volynets, S. Lee, G. Kim, B. Lee, S. K. Nam, and K. Han, *J. Vac. Sci. Technol., A* **35**, 061310 (2017).
- ²⁰M. J. Kushner, *J. Phys. D: Appl. Phys.* **42**, 194013 (2009).
- ²¹D. Zhang and M. J. Kushner, *J. Vac. Sci. Technol., A* **19**, 524 (2001).
- ²²B. E. E. Kastenmeier, P. J. Matsuo, G. S. Oehrlein, R. E. Ellefson, and L. C. Frees, *J. Vac. Sci. Technol., A* **19**, 25 (2001).
- ²³S. Huang, V. Volynets, J. R. Hamilton, S. Lee, I.-C. Song, S. Lu, J. Tennyson, and M. J. Kushner, *J. Vac. Sci. Technol., A* **35**, 031302 (2017).
- ²⁴J. Tennyson, *Phys. Rep.* **491**, 29 (2010).
- ²⁵J. R. Hamilton, J. Tennyson, S. Huang, and M. J. Kushner, *Plasma Sources Sci. Technol.* **26**, 065010 (2017).
- ²⁶N. Ruckhaberle, L. Lehmann, S. Matejcek, E. Illenberger, Y. Bouteiller, V. Periquet, L. Museur, C. Desfrancois, and J.-P. Schermann, *J. Phys. Chem. A* **101**, 9942 (1997).
- ²⁷M. Bettendorff and S. D. Peyerimhoff, *Chem. Phys.* **99**, 55 (1985).
- ²⁸S.-H. Song and M. J. Kushner, *Plasma Sources Sci. Technol.* **21**, 055028 (2012).
- ²⁹S. Banna *et al.*, *IEEE Trans. Plasma Sci.* **37**, 1730 (2009).
- ³⁰P. J. Matsuo, B. E. E. Kastenmeier, G. S. Oehrlein, and J. G. Langan, *J. Vac. Sci. Technol., A* **17**, 2431 (1999).