Electrically-inactive phosphorus re-distribution during low temperature annealing

Ana Peral,1,2,a) Amanda Youssef,2 Amir Dastgheib-Shirazi,3 Austin Akey,2 Ian Marius Peters,2 Giso Hahn,3 Tonio Buonassisi,2 and Carlos del Caño1

1Universidad Politécnica de Madrid, E.T.S.I. de Telecomunicación, Instituto de Energía Solar, 28040 Madrid, Spain
2Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
3Department of Physics, University of Konstanz, 78457 Konstanz, Germany

An increased total dose of phosphorus (P dose) in the first 40 nm of a phosphorus diffused emitter has been measured after Low Temperature Annealing (LTA) at 700 °C using the Glow Discharge Optical Emission Spectrometry technique. This evidence has been observed in three versions of the same emitter containing different amounts of initial phosphorus. A stepwise chemical etching of a diffused phosphorus emitter has been carried out to prepare the three types of samples. The total P dose in the first 40 nm increases during annealing by 1.4 × 10¹⁵ cm⁻² for the sample with the highly doped emitter, by 0.8 × 10¹⁵ cm⁻² in the middle-doped emitter, and by 0.5 × 10¹⁵ cm⁻² in the lowest-doped emitter. The presence of surface dislocations in the first few nanometers of the phosphorus emitter might play a role as preferential sites of local phosphorus gettering in phosphorus redistribution, because the phosphorus gettering to the first 40 nm is lower when this region is etched stepwise. This total increase in phosphorus takes place even though the calculated electrically active phosphorus concentration shows a reduction, and the measured sheet resistance shows an increase after annealing at a low temperature. The reduced electrically active P dose is around 0.6 × 10¹⁵ cm⁻² for all the emitters. This can be explained with phosphorus-atoms diffusing towards the surface during annealing, occupying electrically inactive configurations. An atomic-scale visual local analysis is carried out with needle-shaped samples of tens of nm in diameter containing a region of the highly doped emitter before and after LTA using Atom Probe Tomography, showing phosphorus precipitates of 10 nm and less before annealing and an increased density of larger precipitates after annealing (25 nm and less).

I. INTRODUCTION

Phosphorus atoms in a silicon lattice can adopt different chemical configurations. P atoms in some configurations donate charge carriers in the conduction band and are considered electrically active. On the other hand, P atoms in other configurations are electrically inactive.

The standard configuration of phosphorus in the silicon lattice is when a substitutional phosphorus atom with five valence electrons replaces a silicon atom. An electron is donated to the lattice in the conduction band, and the P atom remains as a positive-charged fixed ion. The phosphorus atom is called a donor.

Substitutional P used to be considered an electrically active P, and the electron density (n) was assumed to be equal to the substitutional P density. However, in 2006, Altermatt et al. modeled the incomplete ionization of the substitutional P, an effect that up to that moment was not usually taken into account at room temperature, although it was considered an important effect at low temperatures. They verified that at room temperature, for substitutional doping densities of around 1 × 10¹⁸ cm⁻³, up to 25% of substitutional phosphorus atoms are non-ionized, and it was assumed that only the ionized substitutional atoms donated free electrons. For concentrations above 3 × 10¹⁹ cm⁻³, all of the substitutional P is ionized, so that the density of the substitutional P equals the free electron density.

The explanation is that when the Fermi level is situated close to the dopant energy level in the bandgap, the dopant states are occupied noticeably, leading to an incomplete ionization of the dopants. The free carrier density is then noticeably smaller than the substitutional dopant density, although all the dopant atoms replace the Si atoms.

Since the free charge carrier density can be easily measured, the maximum carrier concentration, n, has been parametrized empirically as a function of the process temperature in accordance with Eq. (1), which has been proposed by Solmi et al. in a work that was further elaborated in Ref. 4.

\[
  n_{\text{max}} [\text{cm}^{-3}] = 1.3 \cdot 10^{22} \cdot \exp (-0.37/k \cdot T),
\]

where \(k\) denotes the Boltzmann constant (equal to 8.617 × 10⁻³⁵ eV/K) and \(T\) is the temperature in Kelvin. Bentzen et al. also developed a similar charge carrier density parametrization.

Most studies into the behavior of phosphorus with temperature have been implemented at high temperature (i.e., temperatures above 800 °C). These studies usually only focus on the variation in the electrically active phosphorus.

The difference between the total chemical P concentration and charge carrier density is usually attributed to P

\[a)\text{Electronic mail: ana.peral@ies-def.upm.es}\]
atoms forming complexes that are electrically inactive. Several deactivation mechanisms have been proposed for phosphorus in silicon. There are observations of the formation of large dopant precipitates as Si-P phases,\(^3\)\(^4\)\(^5\) P clusters, or rod-like P precipitated structures or small defect clusters comprising only a few dopants and native point defects. Some authors attribute the dominating P deactivation effect to phosphorus-vacancy clusters (\(P_n V\) for \(1 \leq n \leq 4\)).\(^6\) Other deactivation proposals are attributed to mobile species containing P atoms, such as interstitial P or interstitial Si-P pairs.\(^7\) Calculations of some of the configurations studied for P atoms, including several binding energy values, can be found in Ref.9.

Solmi et al. parameterized values of the P concentration in Si in balance with large monocrystalline (or orthorhombic) SiP precipitates as a function of temperature.\(^3\) This P concentration equals the sum of the substitutional P and other mobile electrically inactive configurations of P. Bentzen et al. also showed an empirical expression of solid solubility of P in Si,\(^10\) which is in agreement with the solid solubility calculated by Solmi et al.\(^3\)

Several authors described a local accumulation of inactive and immobile P atoms on the surface of the wafer.\(^7\)\(^11\) They found enhanced nucleation at the interface of suitably oriented crystals, adjacent to the doped oxide phosphosilicate glass (PSG) layer. They observed that kind of aggregation after high isothermal temperature predeposition processes above 900 °C without cooling. However, other authors state that temperatures above 600 °C are high enough for deactivating a large amount of the dopant in the first seconds of the annealing, based on the high diffusivity of dopants and point defects.\(^12\)

In addition to large second-phase precipitates, small particles (with diameters of between 1.5 and 7.5 nm) were detected in surface Si grains and deeper in the substrate (also at concentrations below the solid solubility of P in silicon) using weak-beam, dark-field images, and they were suggested to be related to P, but with no final conclusions.\(^3\)

Inactive mobile dopants have been described as an aggregation phenomenon with fluctuant heterophases, which precedes the transition to immobile phases present in balance.\(^3\) It has been suggested that this composition accounts for the behavior of the dopant diffusion and the electrical inactivity of the dopant.

The total P profile distribution has been observed to evolve with several factors. An increased amount of electrically inactive P has been observed with an increasing POCl\(_3\) concentration and with a decreasing process temperature.\(^6\)\(^7\) Furthermore, an increased size of P precipitates has been observed by increasing the process time.\(^11\) The SiP precipitates region below the surface was observed to shift toward the surface in the P concentration profile, when increasing the annealing time for different temperatures (800, 850, and 900 °C). In addition, P was observed to diffuse both toward the growing SiP precipitates and to the silicon bulk.\(^3\)

Understanding the evolution of the electrically active and the electrically inactive fractions of P with thermal processes is relevant, as they impact the recombination within the emitter, and can help in the design of techniques for extrinsic recombination reduction as a result of defects in the emitter. In particular, the effect of using a low-temperature annealing after phosphorus diffusion on the emitter recombination in solar cells has been measured, leading to the process parameters that reduce Shockley–Read–Hall (SRH) recombination and hence improve solar cell performance.\(^13\)

Note that by low temperature, we refer to temperatures below the usual POCl\(_3\) deposition temperatures, i.e., temperatures below 800 °C.

The aim of this work is to understand further the effect of the low temperature annealing (LTA) on P distribution, observing the changes in P distribution and the influence of the initial electrically inactive content of the P profile on final P distribution.

II. MATERIAL AND METHODS

B-doped (30 Ω cm) single-side polished Czochralski-grown Si wafers of 300 μm thickness received a standard chemical cleaning, an HCl solution dip followed by an HF solution dip, and a P doping using a tubular POCl\(_3\) furnace at a temperature of 875 °C for 1 h of pre-deposition and 30 min of drive-in. After diffusion, the PSG was etched using 2 min of HF solution (1%). A sheet resistance of 17 Ω/√\(\square\) was measured. Then, we used an acidic wet-chemical emitter etch-back (EEB) procedure,\(^14\) to obtain three sample types (see Fig. 1):

- **EEB 0 nm sample** had a highly doped P emitter containing a high concentration of electrically inactive phosphorus. The measured sheet resistance was 17 Ω/√\(\square\).
- **EEB 20 nm sample** received 2 min of acidic chemical solution, which resulted in etching 20 nm of the surface of the wafer which was expected to reduce the amount of electrically inactive phosphorus content in the emitter. A sheet resistance of 20 Ω/√\(\square\) was measured after etching.
- **EEB 40 nm sample** received 4 min of acidic chemical solution for etching around 40 nm of the surface of the wafer, to obtain an emitter with lower amount of inactive P. A sheet resistance of 21 Ω/√\(\square\) was measured after etching.

The three groups of samples were annealed at 700 °C for 30 min in nitrogen.

![FIG. 1. Distribution of samples.](image-url)
The effect of low temperature annealing on the distribution and configuration of phosphorus was analyzed. Pre- and post-annealing references were measured and compared. Sheet resistance related to the electrically active P content is measured by means of a four point probe. The total chemical P concentration was measured using Glow Discharge Optical Emission Spectrometry (GD-OES). In the GD-OES, the energy released to the sample originates from ions from a glow discharge, a special case of plasma formed by the passage of an electric current through a low-pressure gas. The analysis area provided is around 2 mm², and the minimum depth of information is 0.1 nm. This technique allows the analysis of the total P concentration in silicon and in SiO₂ layers, without providing information about the chemical or electrical configurations of the phosphorus atoms. Atom Probe Tomography (APT) images are included for the spatial analysis of the size and shape of the P precipitates and also for measuring their distribution inside the kink part of the P profile. Needle-shaped samples are prepared from both the pre-annealed and the post-annealed samples using focused-ion-beam lift-out in a Zeiss NVision 40 Dual Beam system for APT measurements. A protective layer of platinum, a few hundred nanometers thick, is deposited by the ion beam. A high voltage pulse is applied to a specimen, previously prepared with needle geometry using focused ion beam methods as the standard lift-out procedure. Laser pulsed APT measurements were taken using a LEAP 4000X HR with a 20 pJ laser pulse energy, a 100 KHz pulse rate, and at a base temperature of 40 K. Atoms, progressively field evaporated from the tips, are projected onto a position-sensitive detector and identified by time-of-flight mass spectrometry. We have reconstructed the tips in 3D by following the geometry of the tips in the scanning electron micrographs, assuming a spherical tip surface. In this case, 3D atomistic information is provided, and then, the analysis area is the size of the needle (around tens of nm), and the minimum depth information is 0.1 nm. Both techniques can only be compared qualitatively because the size of the measured sample is different.

III. RESULTS AND DISCUSSION

We measured the total P density profile of the different types of emitter. Fig. 2 shows the depth profile of the 3 types of P emitter measured using the GD-OES technique before and after low-temperature annealing (black and red lines, respectively). The region with a concentration higher than $1 \times 10^{20} \text{cm}^{-3}$ is usually called the “dead layer” because of its high P concentration and content of P atoms in non-electrically active configurations. Each figure corresponds to a different emitter type depending on its content of electrically inactive P. Dashed lines correspond to calculated electron concentration values using Eq. (1) by Solmi et al. To calculate the electron concentration before annealing (black dashed line), the selected temperature was the emitter formation temperature (875 °C), resulting in a free-electron concentration $n = 3.3 \times 10^{20} \text{cm}^{-3}$. To calculate the post annealed electron concentration (red dashed line), we used the annealing temperature (700 °C), obtaining $n = 1.8 \times 10^{20} \text{cm}^{-3}$. According to the calculations, electron density and substitutional phosphorus are expected to be reduced during annealing.

For a quantitative analysis, we calculated the P doses contained in the first 40 nm of each wafer and the results are shown in Table I. The first observation is that the total P

![FIG. 2. Total phosphorus measured using the GD-OES technique before and after annealing for three different P emitters. Dashed lines correspond to calculated electron concentration $n$ using Eq. (1) by Solmi et al.](image)
TABLE I. Effect of LTA on the total P concentration distribution measured by GD-OES in the first 40 nm. The acronyms used in the table are pre = pre-annealing, post = post-annealing, Abs = Absolute, Rel = Relative, Var = Variation, and EEB = Emitter etch-back.

<table>
<thead>
<tr>
<th></th>
<th>P dose pre (\times 10^{15} \text{ cm}^{-2})</th>
<th>P dose post (\times 10^{15} \text{ cm}^{-2})</th>
<th>Abs P dose var (\times 10^{15} \text{ cm}^{-2})</th>
<th>Rel var</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEB-0</td>
<td>3.3</td>
<td>4.7</td>
<td>1.4</td>
<td>0.41</td>
</tr>
<tr>
<td>EEB-20</td>
<td>1.7</td>
<td>2.5</td>
<td>0.8</td>
<td>0.49</td>
</tr>
<tr>
<td>EEB-40</td>
<td>1.3</td>
<td>1.8</td>
<td>0.5</td>
<td>0.38</td>
</tr>
<tr>
<td>Electrically inactive P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEB-0</td>
<td>2</td>
<td>3.9</td>
<td>1.9</td>
<td>0.97</td>
</tr>
<tr>
<td>EEB-20</td>
<td>0.4</td>
<td>1.8</td>
<td>1.4</td>
<td>3.79</td>
</tr>
<tr>
<td>EEB-40</td>
<td>0.07</td>
<td>1.1</td>
<td>1.0</td>
<td>15.10</td>
</tr>
<tr>
<td>Electrically active P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEB-0</td>
<td>1.3</td>
<td>0.7</td>
<td>−0.60</td>
<td>−0.45</td>
</tr>
<tr>
<td>EEB-20</td>
<td>1.3</td>
<td>0.7</td>
<td>−0.60</td>
<td>−0.45</td>
</tr>
<tr>
<td>EEB-40</td>
<td>1.26</td>
<td>0.7</td>
<td>−0.56</td>
<td>−0.43</td>
</tr>
</tbody>
</table>

Dose in the first 40 nm increases during annealing in the three emitter types: An extra P dose of \(1.4 \times 10^{15} \text{ cm}^{-2}\) in the case of highly doped emitter, \(0.8 \times 10^{15} \text{ cm}^{-2}\) in the case of middle-doped emitter, and \(0.5 \times 10^{15} \text{ cm}^{-2}\) in the lowest-doped emitter. This represents a relative variation similar in all the emitters of \(+40\%–50\%\), compared to the initial total P dose. According to this result, the increase in P within the first 40 nm has a correlation with the initial total P dose. This effect could be related to the dislocations present in the first few nanometers of the wafer or to the P precipitates that also induce defects in the crystal, that have been etched stepwise in the case of wafers from groups 2 and 3. This kind of dislocation has been observed as the preferential sites of a local accumulation of phosphorus by other authors.\(^7\) This experiment only can confirm the increase in P; however, the source of P for this increase cannot be confirmed and is part of future research. An external source of P was discarded with extra control bare wafers placed in the furnace during the same annealing. They were measured using an Electrochemical Capacitance-Voltage profiler without measuring any detectable incorporation of phosphorus from the furnace. These measurements were also confirmed using a four point probe and a hot-spot probe. Phosphorus at low concentrations is expected to occupy the electrically active positions first, so that measuring a zero electrically active P concentration should be equal to having a zero total P concentration. Phosphorus redistribution from a deep region of more than 130 nm from the surface should be studied in future experiments.

Calculations regarding electrically active P present using Eq. (1) by Solmi et al. with the P diffusion temperature of 875°C shown in Table I reveal a dose value of around \(1.3 \times 10^{15} \text{ cm}^{-2}\) in the surface 40-nm region for all the 3 emitter types before the annealing. After the annealing at 700°C, this electrically active P dose is reduced to \(0.7 \times 10^{15} \text{ cm}^{-2}\) in the same 40-nm region. Sheet resistance measurements show an increase of 3 \(\Omega \cdot \square\) after LTA which is coherent with the calculated reduction in the P dose. As expected, the calculated dose of electrically inactive P is increased \(+97\%\) in the highly doped emitter (corresponding to an absolute dose of \(1.9 \times 10^{15} \text{ cm}^{-2}\)), \(+379\%\) in the middle-doped emitter (corresponding to \(1.42 \times 10^{15} \text{ cm}^{-2}\)), and \(+1510\%\) (corresponding to an extra dose of \(1.03 \times 10^{15} \text{ cm}^{-2}\)).

We observe that the absolute dose increase is similar in the surface 40-nm for the three emitter types, and a direct correlation between the increase in the dose and initial dose is present. As can be observed during annealing at low temperature, the P atoms diffuse toward the surface region and they occupy an electrically inactive configuration. This idea is consistent with the growing P precipitates theory postulated by other authors for high-temperature processes.\(^5\),\(^6\),\(^11\),\(^16\)

The molar concentration percentage of phosphorus, oxygen and silicon of an EEB-0 nm sample is represented in Fig. 3. The increase in P in the first 10 nm is observed in detail. There is no correlation in the maximum of P with oxygen or silicon total concentration. Other authors detected similar agglomerations attributing them to a different crystallinity region in the first nm of the surface.\(^3\)

A sample without surface etching (EEB-0 nm sample) was measured with APT before and after annealing. APT data showing 3 percent iso-surface of the phosphorus in the emitter layer are represented in Fig. 4, and 10% iso-surface in the same sample can be observed in Fig. 5. The dimensions in the APT tip are in nanometers.

Precipitates in the order of 10 nm and less are seen locally in the APT measurement of the pre-annealed sample [Figs. 4(a) and 5(a)]. The majority of P clusters are allocated at a distance of 10 nm of the surface and deeper.

In the post-annealed sample, we observe larger precipitates locally in the order of 20 nm and less, as well as a higher precipitate density. A large precipitate 25 nm in diameter in the surface can be observed in Fig. 4(b) and in Fig. 5(b). Other authors also observed (for high-temperature annealings) larger sized phosphorus precipitates after annealing compared with the measurements before annealing.\(^3\),\(^11\)

APT observations are in qualitative agreement with the observations made in the total P measurement using...
GD-OES, as can be seen representing a molar concentration measured using both measurement techniques in Fig. 6, which are in the same range. Although APT data give only very localized information and cannot be used for a good statistical evaluation, it is nevertheless a very good hint that the observed increase in P dose close to the Si surface after LTA is related to more and larger P precipitates in this region.

IV. CONCLUSIONS

Low temperature annealing has been demonstrated to have an effect on the distribution of phosphorus and its chemical state. Many researchers have studied the behavior of phosphorus at high temperatures, but these studies usually only focus on electrically active phosphorus, whose change at low temperatures is negligible. These changes in P configuration within the emitter at low temperatures can be used as a means of reducing SRH recombination in the emitter, as has been demonstrated in previous works.\textsuperscript{13}

In this work, total phosphorus distribution has been studied during low temperature annealing, observing an increase in P in the surface region (the first 40 nm). There is a correlation in the increase in P with initial P dose present. It has been calculated that electrically active P is reduced during annealing and that the increase in total P is caused by P occupying electrically inactive positions.

Atomic-scale visual local analysis is carried out with samples of tens of nm using APT before and after LTA, showing phosphorus precipitates of 10 nm and less before annealing, and an increased density of larger-sized precipitates after annealing (25 nm and less). The changes observed in the measured area of APT (tens of nm\textsuperscript{2}) are in agreement with the observations of the mean value in the measured area of GD-OES (2 mm\textsuperscript{2}).

This experiment can only confirm the increased overall P concentration after a 30 min annealing at 700 °C in nitrogen. However, the P source for this increase cannot be confirmed yet. An external P source is discarded by an extra experiment, so that P redistribution from a deep region (more than 130 nm from the surface) might be the cause, a hypothesis that needs to be further studied.

ACKNOWLEDGMENTS

Amanda Youssef acknowledges the funding provided by the National Science Foundation (NSF) and the Department
Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of NSF or DOE. This work was carried out in part at the Center for Nanoscale Systems (CNS), a member of the National Nanotechnology Infrastructure Network (NNIN), which was supported by the National Science Foundation under Contract No. ECS-0335765. Ana Peral gratefully acknowledges the Fellowship for Visiting Scientists Stays (EEBB-I-14-08268 and EEBB-I-15-09198) of the Program for Training of Researchers of the Ministry of Economy and Competitiveness of Spain. This work has been partly funded by the Spanish Ministerio de Ciencia e Innovacion through TABACO Project ENE2014-56069-C4-2-R project, and by Comunidad de Madrid through the MADRID-PV S2013/MAE-2780 project.

1S. M. Sze and K. K. Ng, Physics of Semiconductor Devices (John Wiley and Sons, 2006).