

NON-MELT THIN-WAFER LASER THERMAL ANNEALING METHODS

FIELD

[0001] The present disclosure relates generally to laser annealing of semiconductor materials when forming integrated circuit structures, and in particular relates to methods of performing non-melt laser thermal annealing of thin semiconductor wafers.

BACKGROUND ART

[0002] Laser thermal annealing (LTA) is used in semiconductor manufacturing for a variety of applications, including for activating dopants in select regions of devices (structures) formed in a semiconductor wafer when forming active microcircuits such as transistors and related types of semiconductor features.

[0003] One type of laser annealing uses a scanned line image from a light beam to heat the surface of the wafer to a temperature (the “annealing temperature”) for a time long enough to activate the dopants in the semiconductor structures (e.g., source and drain regions) but short enough to prevent substantial dopant diffusion. The time that the wafer surface is at the annealing temperature is determined by the power density of the line image, as well as by the line-image width divided by the velocity at which the line image is scanned (the “scan velocity”). The amount of time the line image resides over a point on the wafer surface is called the “dwell time.”

[0004] For some semiconductor device applications, there is a need to heat one side of the wafer while keeping the other side below a certain critical temperature. One example is a power device, where back-side dopant activation and contact is carried out after the front device is fabricated. Another example is a thin-substrate solar cell, wherein a highly conductive emitter is needed on the front side to improve the cell efficiency, and which can be achieved by doping and annealing. A third example is a back-side image sensor, where a field-stop layer that typically consists of highly doped and activated layers is used on the back side to suppress the dark current.

[0005] In all of these examples, there is a need for thermal annealing to activate dopants or form a contact on one side of the wafer. The substrate thicknesses involved are in the

range of several to a few hundred micrometers (i.e., microns or μm), which is significantly thinner than the thickness of a standard 8-to-12-inch silicon wafer, namely, 725 μm to 775 μm . Typical dopant activation requires an annealing temperature above about 1,000°C. However, the maximum temperature and thermal budget for the other side of the substrate must be limited to avoid any potential degradation in material integrity and junction dopant profiles associated therewith.

[0006] For example, if metallization is already present on a first side of the wafer prior to laser thermal annealing, the maximum temperature of the first side needs to be kept below the melting point of the metal in order to maintain good physical integrity while the opposite (second) side is annealed.

[0007] Conventional rapid thermal annealing (RTA) has an annealing time on the order of seconds, which in silicon corresponds to a thermal diffusion length L_D of several millimeters. This length is significantly greater than the typical wafer thickness, which means that both sides of the wafer will see similar peak annealing temperatures during RTA. Consequently, RTA is not suitable for such thin-wafer applications.

[0008] Conventional laser annealing methods for such thin-wafer applications involve using a pulsed melt laser with a pulse length in the range of tens to hundreds of nanoseconds. For silicon, this corresponds to a thermal diffusion length on the order of about 1 μm , which is significantly less than most wafer thicknesses. An advantage of pulse-melt annealing is that it can locally heat one side of a wafer to a very high temperature (including above the silicon melting temperature) with almost no heat penetration to the other side. Thus, it can be used for silicon wafers as thin as a few micrometers.

[0009] However, in the nanosecond time frame associated with pulse-melt annealing, dopant activation can be achieved only in the melted state. During melting, dopants can quickly diffuse and redistribute into a more box-like profile. While this may be advantageous for some applications, it may not be desirable in cases where precise doping profiles need to be maintained. In particular, the melting of multiple junctions can cause dopants of opposite polarities to intermix and degrade junction performance. A second limitation in the pulsed-melt annealing method is the depth of heat penetration. Because of the short thermal diffusion length, it is difficult to effectively anneal junctions deeper than 0.5 μm .

Also, the method cannot be effectively used to anneal out implant defects beyond the melt depth due to the extremely low thermal budget associated with nanosecond-scale pulses.

[0010] Accordingly, there is a need for annealing thin semiconductor wafers that allows for efficient non-melt dopant activation on one side of the wafer without causing detrimental heating effects on the other side of the wafer.

SUMMARY

[0011] This disclosure is directed to methods of annealing a thin semiconductor wafer such that high-temperature annealing can be performed on one side of a thin semiconductor wafer without damaging or overheating the heat-sensitive features or devices on the other side of the wafer (or embedded within the wafer). The peak temperature is below the melt point of the wafer so that no significant dopant re-distribution occurs during the annealing process. The methods can be applied to activating dopants or forming ohmic contacts.

[0012] An aspect of the disclosure is a method of annealing a semiconductor product wafer having a back side with an annealing region, the product wafer having electronic device features that reside a distance d from the back side and that are subject to being damaged beyond a critical temperature T_C . The method includes scanning an annealing laser beam over the back side to anneal the annealing region by bringing the annealing region up to an anneal temperature T_A that is less than a melt temperature T_M of the semiconductor product wafer. The scanning has a thermal diffusion length L_D attendant therewith, and the method further includes performing the scanning with a dwell time such that a thermal diffusion length L_D satisfies $L_D < d$ and wherein the electronic device features are maintained below the critical temperature T_C .

[0013] Another aspect of the disclosure is the method disclosed above, wherein the annealing region is either an ion-implant layer or a contact layer.

[0014] Another aspect of the disclosure is the method disclosed above, wherein the dwell time is in the range from 3 microseconds to 200 microseconds.

[0015] Another aspect of the disclosure is the method disclosed above, wherein the electronic device features include metal and wherein the critical temperature T_c is in the range from about 600 °C to about 900 °C.

[0016] Another aspect of the disclosure is the method disclosed above, further including forming the semiconductor product wafer by interfacing a device wafer and a carrier wafer, wherein the device wafer includes the electronic device features.

[0017] Another aspect of the disclosure is the method disclosed above, wherein the annealing laser beam has a visible wavelength or an infrared wavelength with an optical absorption length less than the thin wafer thickness.

[0018] Another aspect of the disclosure is the method disclosed above, wherein the electronic device features include a CMOS device layer.

[0019] Another aspect of the disclosure is the method disclosed above, wherein the electronic device features include a semiconductor power-device layer.

[0020] Another aspect of the disclosure is the method disclosed above, further comprising performing multiple scans of a same portion of the back side, where temporally adjacent scans are separated in time by a time interval τ , wherein $1 \text{ millisecond} \leq \tau \leq 10 \text{ seconds}$.

[0021] Another aspect of the disclosure is the method disclosed above, wherein the number of multiple scans is 10 or fewer.

[0022] Another aspect of the disclosure is the method disclosed above, wherein $5 \mu\text{m} \leq d \leq 150 \mu\text{m}$.

[0023] Another aspect of the disclosure is the method disclosed above, wherein $10 \mu\text{m} \leq d \leq 20 \mu\text{m}$, the electrical device features are made of aluminum, and a dwell time t_d is in the range from about 3 microseconds to about 20 microseconds.

[0024] Another aspect of the disclosure is the method disclosed above, wherein $5 \mu\text{m} \leq d \leq 150 \mu\text{m}$, the electrical device features are made of copper, and dwell time t_d is in the range from about 15 microseconds to about 100 microseconds.

[0025] Another aspect of the disclosure is a method of annealing a product wafer. The method includes forming a product wafer by forming a protective structure over a front side of a device wafer having a back side and a thickness d between 5 microns and 150 microns, wherein the front side includes electronic device features that are subject to being damaged if heated beyond a critical temperature T_C . The method includes forming an annealing region in or on the back side of the device wafer. The method also includes scanning an annealing laser beam over the back side to perform non-melt annealing of the annealing region, wherein the scanning is performed with a dwell time t_d that defines a thermal diffusion length $L_D < d$ and wherein the electronic device features are maintained at a temperature below the critical temperature T_C .

[0026] Another aspect of the disclosure is the method disclosed above, wherein the annealing region is either an ion-implanted doped layer formed in the back side or a contact layer formed on the back side.

[0027] Another aspect of the disclosure is the method disclosed above, further includes performing multiple scans of a same portion of the back side, wherein temporally adjacent scans are separated by a time interval τ , and wherein $1 \text{ millisecond} \leq \tau \leq 10 \text{ seconds}$.

[0028] Another aspect of the disclosure is the method disclosed above, wherein the number of multiple scans is 10 or fewer.

[0029] Another aspect of the disclosure is the method disclosed above, wherein the protective structure includes a carrier wafer.

[0030] Another aspect of the disclosure is the method disclosed above, wherein the protective structure consists of a single protective layer.

[0031] Another aspect of the disclosure is the method disclosed above, wherein the electronic device features include at least one of a metal and a CMOS layer.

[0032] Another aspect of the disclosure is the method disclosed above, wherein the electronic device features include a semiconductor power-device layer.

[0033] Another aspect of the disclosure is the method disclosed above, wherein the dwell time t_d is in the range from 3 microseconds to 200 microseconds.

[0034] Another aspect of the disclosure is the method disclosed above, further includes selecting the dwell time t_d to be less than a maximum dwell time t_{dm} defined by $t_{dm} = d^2/D_{eff}$ where D_{eff} is the effective thermal diffusivity of the device wafer.

[0035] Another aspect of the disclosure is the method disclosed above, wherein the annealing laser beam has a visible wavelength or an infrared wavelength with an optical absorption length less than the thickness of the thin wafer.

[0036] Another aspect of the disclosure is the method disclosed above, wherein $10\ \mu\text{m} \leq d \leq 20\ \mu\text{m}$, the electrical device features are made of aluminum, and wherein the dwell time t_d is between about 3 microseconds to about 20 microseconds.

[0037] Another aspect of the disclosure is the method disclosed above, wherein $10\ \mu\text{m} \leq d \leq 20\ \mu\text{m}$, the electrical device features are made of copper, and wherein the dwell time t_d is between about 15 microseconds to about 100 microseconds.

[0038] All references cited herein are incorporated by reference herein.

[0039] The claims as set forth below are incorporated into and constitute a part of the Detailed Description.

[0040] In the discussion presented herein, "d" is used interchangeably for both "distance" and "thickness," as these two concepts are closely related as will be understood by one skilled in the art. In some cases, the distance d is the wafer thickness, while more generally the distance d is the distance as measured from the wafer back side to the electronic device features, which can also be the same as the wafer thickness.

BRIEF DESCRIPTION OF THE DRAWINGS

[0041] FIGS. 1A through 1C are various views of example product wafers that can be subject to laser annealing using the laser annealing systems and methods of the present disclosure;

[0042] FIG. 2A through FIG. 2D are side views of four different example product wafers and include respective close-up cross-sectional views of the example product wafers;

[0043] FIG. 3 is a schematic diagram of an example laser annealing apparatus suitable for carrying out the annealing methods disclosed herein;

[0044] FIG. 4 is a plot of Temperature ($^{\circ}\text{C}$) versus Time that shows the process windows for four different annealing methods, including the annealing methods disclosed herein;

[0045] FIG. 5 is a plot of Temperature vs. Depth into a thin silicon wafer and compares the depth temperature profiles across the product wafer for four different annealing techniques;

[0046] FIGS. 6A and 6B are plots of Boron Concentration (cm^{-3}) versus Depth (nm) illustrating how a melt annealing process causes substantial dopant diffusion and a substantial change in the dopant profile (FIG. 6A) as compared to the non-melt annealing methods disclosure herein (FIG. 6B);

[0047] FIG. 7 is a plot of Temperature versus Time and shows example temperate profiles for a point on the back side of product wafer and at the location of the electronic device features;

[0048] FIG. 8 plots temperature T_B versus dwell time t_d for two product wafers of thickness d_1 and d_2 , where $d_1 < d_2$ and shows how for a given wafer thickness, the dwell time t_d needs to be below a maximum dwell time t_{dm} to ensure that $T_B < T_C$;

[0049] FIG. 9 is a plot of dwell time t_d versus wafer thickness d and shows the limits on the dwell time for a given substrate thickness, as indicated by the shaded area in the plot;

[0050] FIG. 10 is a plot of Temperature versus Time and illustrates a series of N typical temperature profiles for a given point on the wafer associated with performing multiple (N) scans of the annealing laser beam;

[0051] FIG. 11A is a plot of leakage current (arbitrary units, a.u.) versus the number of annealing scans and illustrates a reduction in leakage current with the number of scans, which indicates increasing defect annealing with multiple/successive annealing scans; and

[0052] FIG. 11B is a plot of sheet resistance R_s (ohm/sq.) versus number of annealing scans and illustrates how the sheet resistance decreases with an increasing number of scans, which is indicative of enhanced dopant activation.

DETAILED DESCRIPTION

[0053] Reference is now made in detail to embodiments of the disclosure, examples of which are illustrated in the accompanying drawings. Whenever possible, the same or like

reference numbers and symbols are used throughout the drawings to refer to the same or like parts.

[0054] The fabrication of certain types of semiconductor devices, such as image sensors and high-power devices, involves the use of relatively thin semiconductor wafers. FIG. 1A is an example side view of a product wafer 10 having a top side 21 that defines a top surface 22 and a back side 23 that defines a back surface 24. The product wafer 10 of FIG. 1A can be considered to be upside down, since the “top side” 21 faces downward in the Figure.

[0055] FIG. 1B illustrates the formation of an example product wafer 10 formed by interfacing a device wafer 10a and a carrier wafer 10b. The device wafer 10a has a front side 12a in which electronic device features 34 are formed that is accordingly called the device side. The device wafer 10a also includes a back side 14a that is opposite to the front side (device side) 12a. The carrier wafer 10b has a front side 12b on which an oxide layer 15 is formed. The resulting wafer 10, referred to herein as a product wafer, is formed by interfacing the device side 12a of device wafer 10a with the front side 12b of carrier wafer 10b. The oxide layer 15 serves as a bonding layer that bonds the two wafers 10a, 10b together. Thus, the front side 12b of carrier wafer 10b is also called the bonding side. An example device wafer 10a is formed from a silicon substrate.

[0056] The resulting product wafer 10 is shown in FIG. 1C. At this point, the device wafer 10a is ground down from the back side 14a to reduce the thickness of the device wafer 10a from about 750 μm to a thickness d in the range from about 5 μm to about 150 μm , as shown in the close-up inset of FIG. 1C. The oxide layer 15 and carrier wafer 10b define a protective structure on top side 21 that serves to protect electronic device features 34, which are now embedded within the product wafer 10.

[0057] FIG. 2A is similar to FIG. 1C and includes an inset that shows a more detailed close-up cross-sectional view of an example CMOS device structure 30 of product wafer 10. The example CMOS device structure 30 is in the form of a CMOS sensor. The carrier wafer 10b with the oxide layer 15 (which is typically several microns thick) supports the device wafer 10a, which includes heat-sensitive electronic device features 34. Here, the term “heat-sensitive” means that electronic device features 34 can be damaged if they are heated to beyond a critical temperature T_c .

[0058] An example of heat-sensitive electronic device features 34 includes metallization elements, which are shown in FIG. 2A as being in contact with an adjacent CMOS device layer 40. The CMOS device layer 40 is in turn supported by the thinned device wafer 10a, which has a thickness d in the range from about 5 μm to 50 μm . In another example of a power device, thickness d is in the range from about 30 μm to about 150 μm .

[0059] In an example, a thin ion-implant layer 44 that needs to be laser annealed is formed in the thinned device wafer 10a adjacent its back side 14a, which defines surface 24. The ion-implant layer 44 includes dopants used to form junctions, to reduce leakage current or to lower contact resistance. Example dopants for silicon are arsenic, phosphorus, boron, indium, antimony, etc. Thermal annealing is needed to repair implant damage and to activate the dopants. This process is called back-side-junction activation.

[0060] FIG. 2B is similar to FIG. 2A and includes a contact layer 54 atop ion-implant layer 44. The annealing of contact layer 54 results in the alloying of the contact layer 54 with a silicon substrate to form an ohmic contact and is referred to as back-side contact annealing. The contact layer 54 typically consists of thin-film metal stacks. Example metals used to form the contact layer 54 include Ti, Ni, Co, W, Al, Cu, Au, Pt, etc. Laser annealing of the contact layer 54 can be used to form a metal silicide and to lower the contact resistance of the contact layer 54.

[0061] The annealing of contact layer 54 can also help to activate doped ion-implant layer 44 underneath the contact layer 54. However, the maximum annealing temperature in this case may be limited by the silicide degradation or damage to the metal-film stack, and therefore the dopant activation may not be as optimized as in the configuration of product wafer 10 of FIG. 2A. An alternative is to apply laser annealing twice (before and after deposition of the contact layer 54) to separately optimize the dopant activation and contact formation.

[0062] FIG. 2C is similar to FIG. 2A, except that product wafer 10 is not formed by interfacing device wafer 10a with carrier wafer 10b. Rather, the device wafer 10a is provided with a protective layer 19 that covers electronic device structures 34 adjacent the front side 12a of the device wafer 10a to form the top side 21 of product wafer 10. FIG. 2D is the same as FIG. 2C, except that it includes the contact layer 54 similar to that shown in FIG.

2B. The protective layer 19 is used, for example, to avoid damaging the electronic device features 34 when supporting the product wafer 10 using a chuck 110 (see FIG. 3). Examples of the protective layers 19 include silicon dioxide and nitride. Thus, in an example, a protective structure 17 consists of the protective layer 19.

[0063] In the example CMOS device structure 30, electronic device features 34 of device wafer 10a reside from about 5 μm to about 150 μm away from the back side 14a of device wafer 10a, which is usually not patterned. Thus, in an example, the back side 14a of device wafer 10a defines a flat and unpatterned back surface 24 of the product wafer 10 that makes for a good annealing surface.

[0064] FIGS. 2A through 2D each show an annealing laser beam 80 being scanned over the back side 23 of the product wafer 10 in carrying out the non-melt laser thermal annealing according to the disclosure. The back side 23 of the product wafer 10 includes an annealing region, such as ion-implant layer 44 or contact layer 54, while heat-sensitive electronic device features 34 reside a distance d away from the back side 14a of the device wafer 10a. The annealing region needs be annealed in a manner that does not damage heat-sensitive electronic device features 34 that reside a distance d away, wherein, in an example, d is between 5 μm and 150 μm .

[0065] FIG. 3 illustrates an example laser annealing system 100 that can be used to perform the non-melt laser annealing methods disclosed herein. The example laser annealing systems 100 that can be used to carry out the non-melt laser thermal annealing methods disclosed herein are disclosed in U.S. Provisional Patent Application Serial No. 61/658,086, entitled "Laser annealing systems and methods with ultra-short dwell times." With reference to FIG. 3, the laser annealing system 100 includes a scanning laser system 150 and a chuck 110 that supports the product wafer 10. In an example, the chuck 110 is cooled. Also in an example, the chuck 110 is movable.

[0066] The scanning laser system 150 includes a laser 151 that generates an initial laser beam 80' along a first axis A1 toward a scanning mirror 162 that is operably connected to a mirror driver 164. The scanning laser system 150 includes a focusing lens 166 oriented along a second axis A2 that is generally at a right angle to the first axis A1. The second axis A2 intersects the first axis A1 at the scanning mirror 162 and also intersects the back side 23 of

the product wafer 10 at substantially a right angle. The focusing lens 166 has a numerical aperture NA. In an example, the laser 151 generates infrared light so that annealing laser beam 80 includes an infrared wavelength.

[0067] In another example, the laser 151 generates visible light 80' so that annealing laser beam 80 includes a visible wavelength. A visible wavelength for light 80' is preferred for very thin substrates (e.g., $d < 30\mu\text{m}$) because of the shorter optical absorption depth. In general, it is desirable for the optical absorption depth to be smaller than the thin wafer thickness.

[0068] A collimating lens 168 is shown adjacent the laser 151 and receives and collimates divergent laser light 80' from the laser 151. The collimated laser light 80' is reflected by the scanning mirror 162 to the focusing lens 166, which focuses the collimated laser beam 80' to form focused annealing laser beam 80. The focused annealing laser beam 80 in turn forms a line image 82 on the back side 23 of the product wafer 10. The movement of scanning mirror 162 causes the annealing laser beam 80 and line image 82 to move rapidly over the back side 23 of product wafer 10, as indicated by arrow AR2. In an example, annealing laser beam 80 includes a visible wavelength.

[0069] The distance from the back side 23 of product wafer 10 to the focusing lens 166 is DW. An example distance DW is about 1 meter and an example NA is about 0.15. The mirror driver 164 serves to drive the scanning mirror 162, e.g., to rapidly rotate the scanning mirror 162 through a select angular range so that the annealing laser beam 80 and line image 82 can scan over a corresponding select angular range denoted in FIG. 3 as θ . In an example, the angular range θ is selected so that the line image 82 can be scanned across the back side 23 of product wafer 10 from one edge of the product wafer 10 to the opposite edge at the widest part of the product wafer 10.

[0070] In an example, annealing laser beam 80 simply sweeps from side to side while between scans product wafer 10 is translated in the cross-scan direction (e.g., by moving the chuck 110, as indicated by dashed arrow AR3) so that line image 82 exposes (irradiates) different portions of the back side 23 of product wafer 10 on each scan or at least covers some new portion of the product wafer 10 on adjacent scans (i.e., there can be some overlap of adjacent scans). Generally, the chuck 110 is movable in three dimensions and is

also rotatable in any direction, as needed. In an example, annealing laser beam 80 scans at least a portion of the back side 23 of product wafer 10.

[0071] The annealing temperature is controlled by the laser power and the anneal time. The anneal time is defined by the dwell time, which is the ratio of laser beam (or line image) width to scanning speed. A typical laser beam width ranges from several microns to a few hundred micrometers, and a typical scanning speed ranges from 0.1 to 5 m/sec.

[0072] The annealing conditions of product wafer 10 are such that the device side 12a of product wafer 10 must remain at relatively low temperatures and in particular below the critical temperature T_C , which is associated with the electronic device features 34. The critical temperature T_C is that temperature beyond which the electronic device features 34 will be damaged. Thus, the annealing method is carried out such that electronic device features 34 remain below the critical temperature T_C in order to protect the ultimate functionality of the device. The maximum temperature at device side 12a is determined by the particular electronic device features 34. For metals used for electronic device features 34 such as interconnects, the critical temperature T_C is about 900°C for copper and 600°C for aluminum. For devices containing silicide, the critical temperature T_C may also be limited by the thermal stability of the silicide material.

[0073] The constraint that electronic device features 34 formed at device side 12a need to remain below a critical temperature T_C gives rise to the requirement that the laser thermal anneal time be short enough to prevent the device surface from becoming too hot. This implies that the thermal anneal time (dwell time) must be such that the corresponding thermal diffusion length L_D is less than the thickness d of device wafer 10a (i.e., $L_D < d$).

[0074] FIG. 4 is a plot of Temperature (°C) versus Time and illustrates the temperature-time process windows for different annealing techniques. FIG. 4 includes four different windows denoted A through D. Conventional rapid thermal processing (RTP) is represented by window A and operates at time scales ranging from seconds to minutes at temperatures of up to 1,200°C. The millisecond-laser and flash-lamp annealing methods are represented by window B and have annealing times ranging from hundreds of microseconds to tens of milliseconds and temperatures just short of the silicon melt point of $T_M = 1,412^\circ\text{C}$. The

pulsed laser melt process is represented by window D and operates at time scales of 10 ns to 1 μ s, with a peak temperature above the silicon melt temperature T_M .

[0075] In FIG. 4, window C is associated with the short-dwell-time, sub-melt laser thermal annealing methods disclosed herein, which operate at annealing times that range from several to a few hundreds of microseconds, and by definition operates at temperatures below the silicon melting temperature T_M . Thus, FIG. 4 illustrates how the short-dwell-time, sub-melt laser thermal annealing window C of the thermal annealing temperature-time process space is distinct from the process windows of the other types of conventional annealing. An example range of annealing times for window C is 3 microseconds to 200 microseconds.

[0076] FIG. 5 is a plot of Temperature versus Depth into a thin silicon wafer and compares the depth temperature profiles across a thin silicon wafer for the above four different annealing techniques. The wafer thickness d is in the range of 5 μ m to 200 μ m. For RTA (curve A), the temperature stays flat through the entire substrate due to the large thermal diffusion length L_D associated with the long anneal times (i.e., $L_D \gg d$).

[0077] For millisecond-laser or flash-lamp annealing methods (curve B), the temperature decays with the depth d into the wafer. But the rate of decrease is too slow, and, as a result, the temperature at the other side of the wafer can still exceed the damage threshold temperature T_C (i.e., $L_D > d$).

[0078] For nanosecond pulsed melt laser annealing (curve D), the temperature drops quickly near the exposed surface, and there is no heat penetration to the other side of the wafer. However, the heat penetration is too shallow to anneal junctions deeper than half a micrometer (i.e., $L_D \ll d$).

[0079] For short-dwell-time sub-melt laser annealing as disclosed herein (curve C), the temperature drop is fast enough to keep the unexposed side of the wafer cool while the heat penetration is still large enough to anneal deep junctions (i.e., $L_D < d$).

[0080] FIGS. 6A and 6B are plots of boron concentration (cm^{-3}) versus Depth (nm) into a silicon wafer doped with boron. FIG. 6A includes two curves: curve C0 shows the dopant redistribution with no anneal (i.e., pre-anneal) and curve CM shows the dopant distribution after a laser melt anneal. The junction profile is determined by the melt depth. The melt

anneal curve CM shows a box-like dopant redistribution profile due to the fast dopant diffusion that occurs when the silicon is in the liquid state. The diffusion stops at the liquid-solid boundary, giving rise to the abrupt box-like profile. An advantage of this approach is that provides a greater amount of dopant activation and lower electrical resistance.

However, the disadvantage is that it causes dopants to redistribute within the melt zone and therefore cannot be applied to cases where precise dopant profiles need to be maintained or to multiple junctions where intermixing of different dopants is undesirable. Further, the activation of dopants is limited to the melt region, and therefore dopants outside the melt region are not effectively activated.

[0081] FIG. 6B includes curve C0 as well as a curve CS associated with the short-dwell-time non-melt laser annealing performed according to the present disclosure. It is observed from FIG. 6B that there is no substantial dopant redistribution due to the short dwell-time non-melt laser annealing. This is because the dopants diffuse at a much slower rate when the silicon wafer remains in the solid state. In addition, deep junctions can also be activated because of the increased thermal diffusion length.

[0082] FIG. 7 is a plot of Temperature versus Time and shows a typical temporal temperature profile associated with short-dwell-time non-melt laser annealing. For dopant activation in silicon, the typical peak annealing temperature T_A at the top surface 22 of the product wafer 10 ranges from 950°C to 1,350°C, as shown in a curve CA. The dwell time is selected to ensure that the maximum bottom temperature T_B at the location of (buried) electronic device features 34 is kept below the damage threshold (i.e., critical) temperature T_C , as shown in a curve CB. For metal silicide contact applications, a lower annealing temperature T_A may be used, e.g., 600°C to 1,000°C.

[0083] FIG. 8 plots the bottom temperature T_B versus dwell time t_d for two wafers of thicknesses d_1 and d_2 , where $d_1 < d_2$. For a given wafer thickness, the dwell time t_d needs to be below a critical or maximum dwell time t_{dm} to ensure that $T_B < T_C$. The maximum dwell time t_{dm} will depend on the wafer thickness. The thinner the wafer, the smaller the necessary dwell time. For the curve associated with wafer thickness d_1 , a maximum dwell time t_{dm1} is shown, and for the curve associated with wafer thickness d_2 , a maximum dwell time t_{dm2} is shown.

[0084] FIG. 9 plots the dwell time versus wafer thickness d based on simulations. The relationship roughly follows a square law, which is consistent with the fact that thermal diffusion length scales as the square root of time. The shaded area of the plot corresponds to the region of the t_d - d process window where device damage would occur. Simulations show that for a wafer thickness of $d = 40\ \mu\text{m}$, a dwell time t_d of less than about $70\ \mu\text{s}$ is needed, assuming the critical temperature T_c is limited by damage to the aluminum metallization. Likewise, for a wafer thickness of $d = 20\ \mu\text{m}$, the dwell time t_d needs to be less than about $20\ \mu\text{s}$.

[0085] In an example where d ranges from $10\ \mu\text{m}$ to $20\ \mu\text{m}$, an example dwell time for aluminum electronic device features 34 is in the range from about 3 microseconds to about 20 microseconds, and for copper electronic device features 34 from about 15 microseconds to about 100 microseconds.

[0086] To the lowest order of approximation, the maximum dwell time t_{dm} can be estimated by

$$t_{dm} \approx \eta \cdot d^2 / D_{\text{eff}}$$

where d is the wafer thickness, D_{eff} is the effective thermal diffusivity of the substrate film stack, and η is a function of damage temperature and is usually on the order of 1. In an example embodiment where damage to aluminum metallization is the limiting factor, the maximum dwell time is approximated by $t_{dm} = d^2 / D_{\text{eff}}$ and the dwell time t_d is selected to be less than this approximate dwell time. In cases where damage to copper metallization is the limiting factor, the maximum dwell time t_{dm} can be much relaxed, typically by a factor of 5.

[0087] One of the key requirements for junction annealing is damage removal capability, which is related to the total thermal budget of the annealing process. In general, the larger the thermal budget, the better the damage removal capability. However, while the short-dwell-time sub-melt laser annealing disclosed herein can avoid overheating electronic device features 34, the reduced dwell time also leads to a smaller thermal budget, which can compromise defect annealing.

[0088] Thus, an example annealing method includes either scanning the same area of back side 23 of product wafer 10 multiple times or increasing the overlap of line image 82

between adjacent scans so that each point on the back side 23 of product wafer 10 is scanned multiple times.

[0089] FIG. 10 is a plot of Temperature versus Time and shows the typical temperature profile that a given point on the back side 23 of product wafer 10 will experience during multiple (N) scans of annealing laser beam 80. The time interval τ between two successive scans should be long enough for the residual heat from the first scan to fully dissipate before the second scan starts. An example range for heat dissipation time τ is 1 millisecond $\leq \tau \leq 10$ seconds, while in a more limited example τ ranges from several milliseconds to a few hundreds of milliseconds.

[0090] FIG. 11A plots the measured junction leakage current (arbitrary units, a.u.) versus the number of scans of annealing laser beam 80. The leakage is a measure of the defect level in the junction depletion region. As the scan number increases, the junction leakage decreases, indicating an improvement in defect annealing.

[0091] An added benefit of using multiple sub-melt laser scans is enhanced dopant activation level. This can be seen from FIG. 11B, which plots the measured sheet resistance R_s (ohm/sq.) versus the number of scans of annealing laser beam 80. A reduction in electrical sheet resistance is observed when the scan number increases, indicating enhanced dopant activation.

[0092] It is noted that a single, long-dwell-time, sub-melt anneal can also provide good activation and has improved defect removing capability. However, the long thermal diffusion length L_D can lead to damage of electronic device features 34. Using multiple short-dwell-time scans offers an increased thermal budget as compared to a single scan approach, while maintaining a thermal diffusion length L_D that is smaller than the wafer thickness d . In an example of the multiple short-dwell-time scan approach, the number of scans is 10 or fewer.

[0093] In principle, one can also employ nanosecond pulsed laser annealing in a sub-melt regime and use multiple pulses to improve the dopant activation. However, because of the extremely short anneal duration associated with each laser pulse, a large number of pulses would be needed, i.e., between about 100 and 1,000 pulses. This approach is relatively slow, which makes it problematic to implement in a mass-production environment. In addition,

the heat penetration depth of the nanosecond pulsed laser annealing is still limited to about 1 μm (as determined by the individual pulse duration) regardless of the number of pulses used, so that the approach is ineffective in activating deep junctions.

[0094] Other factors that could impact the temperature of electronic device features 34 include the temperature of chuck 110 and wafer-to-chuck thermal contact resistance. In general, the lower the chuck temperature or thermal contact resistance, the lower the temperature of electronic device features 34, and hence the longer the dwell time or the higher the top-side annealing temperature can be.

[0095] That said, if the chuck temperature is too low, high thermal stress may result due to the large temperature difference between the top and back sides 21 and 23 of product wafer 10. This situation is undesirable because it can cause slip generation and wafer warpage or even wafer breakage. An example chuck 110 is configured with both heating and cooling functionality so that the chuck temperature can be optimized to balance the need of electronic device features 34 for low stress with the features' need for low temperature. To more effectively clamp the wafer front-side temperature, good thermal contact between product wafer 10 and chuck 110 is desirable. This can be achieved by using a vacuum or electrostatic means. To improve the heat-sink capability, materials with high thermal conductivity can be used for the chuck 110. Example materials for the chuck 110 are aluminum and SiC.

[0096] Although the short-dwell-time, non-melt annealing methods disclosed herein are eminently suited for annealing thin wafers, the methods can be used with structures where surface annealing is needed and where heat-sensitive electronic device features 34 are embedded (buried) a distance d from the back side 23, i.e., the side of product wafer 10 on which the annealing laser beam is incident. In examples of the annealing method, multiple sub-melt laser scans can be used to enhance dopant activation or improve defect annealing without damaging heat-sensitive electronic device features 34.

[0097] It will be apparent to those skilled in the art that various modifications and variations can be made to the present disclosure without departing from the spirit and scope of the disclosure. Thus it is intended that the present disclosure cover the

modifications and variations of this disclosure, provided they come within the scope of the appended claims and their equivalents.

What is claimed is:

1. A method of annealing a semiconductor product wafer formed by interfacing a device wafer and a carrier wafer, wherein the device wafer has a thickness d between a front side having electronic device features and a back side having an annealing region, wherein the electronic device features are subject to being damaged beyond a critical temperature T_C , the method comprising of:

scanning a single annealing laser beam over the back side of the device wafer to anneal the annealing region by bringing the annealing region up to an anneal temperature T_A that is less than a melt temperature T_M of the semiconductor product wafer; and

wherein the scanning of the single annealing laser beam has a thermal diffusion length L_D attendant therewith, and further including performing said scanning by selecting a dwell time t_d that is less than a maximum dwell time t_{dm} defined by $t_{dm} = d^2/D_{eff}$, where D_{eff} is an effective thermal diffusivity of the device wafer so that the thermal diffusion length L_D satisfies $L_D < d$ and the electronic device features are maintained below the critical temperature T_C , wherein $600^\circ\text{C} \leq T_C \leq 900^\circ\text{C}$.

2. The method according to claim 1, wherein the annealing region either includes an ion-implant layer or a contact layer.

3. The method according to claim 1, wherein the single annealing laser beam has a visible wavelength or an infrared wavelength with an optical absorption length less than the thickness of the device wafer.

4. The method according to claim 1, wherein the electronic device features include a CMOS device layer.

5. The method according to claim 1, wherein the electronic device features include a semiconductor power-device layer.

6. The method according to claim 1, wherein said scanning includes performing multiple scans of the single annealing laser beam over a same portion of the back side, wherein temporally adjacent scans are separated in time by a time interval τ , and wherein $1 \text{ millisecond} \leq \tau \leq 10 \text{ seconds}$.

7. The method according to claim 6, wherein the number of multiple scans is 10 or fewer.

8. The method according to claim 1, wherein $10 \text{ } \mu\text{m} \leq d \leq 20 \text{ } \mu\text{m}$, the electrical device features are made of aluminum.

9. The method according to claim 1, wherein $10 \text{ } \mu\text{m} \leq d \leq 20 \text{ } \mu\text{m}$, the electrical device features are made of copper, and wherein the dwell time t_d is between 15 microseconds to 100 microseconds.