

Method and Device for Equalizing a Substrate Stack

Specification

This invention relates to a method for equalizing the thickness variation in a post-bond process of a wafer stack and a corresponding device.

In the modern semiconductor industry the reduction of the layer thickness of a substrate is an essential prerequisite for the continuously advancing miniaturization of the semiconductor components. If a few years ago active and passive components were produced only on a single substrate, especially a wafer, today substrates of different functionality are aligned to one another and bonded to one another. Making contact between the individual function units takes place preferably via through silicon vias (TSVs). The TSVs allow direct production of an electrically conductive connection between the substrates.

One of the critical processes in the production of such a multisubstrate stack is the back-thinning process. In a back-thinning process a substrate with a diameter of generally several hundred millimeters is thinned to a layer thickness of down to less than 50 μm . The substrate is therefore mechanically stabilized during the back-thinning process using a carrier substrate. Moreover the

laser interferometer thickness changes into the nanometer range can be detected and ascertained. A measurement of the substrate stack thickness along the measurement direction takes place by the relative movement between the bonded substrate stack and the laser interferometer. Layer thickness maps and the maximum and minimum layer thickness, and thus the TTV, can be determined from the profile which has been obtained in this way.

The knowledge of the TTV is important mainly for those processes which are applied to the product substrate after bonding, mainly back-thinning and etching processes. An overly large TTV provides specifically for nonuniform thickness removal and thus leads to destruction of the product substrate in the worst case.

Embodiments of the invention disclose a process and a device with which the thickness variation can be reduced, and therefore improved.

According to a first aspect of the present invention, there is provided a method for equalizing the thickness variation in a post-bond process of a wafer stack which consists of a product wafer and a carrier wafer connected by means of an interconnect layer, by local mechanical pressure application of local thickness maxima by means of an application apparatus which has at least one application unit, wherein X-Y-positions of the thickness maxima are determined, wherein the thickness maxima are equalized in a controlled manner by the at least one application unit, wherein the product wafer is disposed to local deformations by the at least one application unit, wherein the at least one application unit applies directly on the surface of the product wafer.

According to a second aspect of the present invention, there is provided a device for equalizing the thickness variation in a post-bond process of a wafer stack which consists of a product wafer and a carrier wafer, which is connected by means of an interconnect layer, with an application apparatus which has at least one application unit for local mechanical pressure application of local thickness maxima, wherein X-Y-positions of the thickness maxima are determined by a measuring device wherein the maxima are equalized in a controlled manner by the at least one application unit, wherein the product wafer is disposed to local deformations by the at least one application unit, wherein at least one application unit applies directly on the surface of the product wafer.

All combinations of at least two of the features given in the specification, the claims and/or the figures also fall within the scope of the invention. At the given value ranges, values within the indicated limits will also be considered to be disclosed as boundary values and will be claimed in any combination. To the extent features disclosed for the device can also be construed as method features, they are also to apply according to the method and vice versa.

The basic idea of embodiments of this invention is to equalize the thickness variation of a substrate stack which is connected by means of an interconnect layer, of a first substrate, in particular a carrier substrate, and a second substrate, in particular a product substrate,

by local application on local thickness extremes, especially thickness peaks, by means of an application apparatus which has at least one application unit. A thickness extreme can be either a thickness peak or a thickness valley. Hereinafter in this patent the application of thickness peaks which is preferred as per embodiments of the invention is used by way of example, but not exclusively, to describe the embodiments of the invention.

A product substrate which is made in particular as a product wafer is thinner than the carrier substrate which is made in particular as a carrier wafer and which is intended for support of the product substrate.

Equalizing within the scope of this invention means in particular that the thickness variation is reduced by the execution of the method as claimed in the invention by 25%, preferably by 50%, still more preferably by at least 75%, most preferably by at least 90%, most preferably of all by at least 95%,

Embodiments of the invention relate especially to a method and a system for reducing the total thickness variation of a bonded substrate stack. One embodiment of the invention lies in particular in a measuring device for determining the TTV or the X-Y position of local peaks, especially in combination with a device which can apply preferably locally on the substrate stack, especially the boundary surface between the two substrates or an interconnect layer between the substrates in order to change thickness fluctuations in a controlled manner, in particular to equalize them. The method as per embodiments of the invention is preferably carried out as a post-bond process. The method as per embodiments of the invention and the system as per embodiments of the invention are preferably used to reduce the TTV between two substrates and a layer in between, in particular a layer of cement. Alternatively

the device and the method can also be used to directly manipulate the TTV of a layer, in particular a layer of cement. In particular therefore all the indicated embodiments and processes as per embodiments of the invention with respect to the direct manipulation of a layer, in particular a layer of cement, are likewise considered to be disclosed. Embodiments of the invention relate in particular a method and a system for reducing, in particular for completely eliminating the in particular total thickness variation of a substrate stack. The substrate stack is preferably a substrate stack which has been produced by a temporary bonding process. The method as per embodiments of the invention can also be fundamentally used for local compensation of unevenness, and thus of the TTV on permanently bonded substrate stacks and/or substrate stacks which have been produced by a prebond.

The fundamental advantage of the embodiment of the invention and of the system as per embodiments of the invention consists in the possibility of a post-bond reduction of the TTV. In terms of process technology the effort will always be made to reduce the TTV completely during bonding. But experience shows that even with the most efficient and most modern bonding systems, a complete reduction of the TTV is not possible.

The system as per embodiments of the invention consists in particular of at least two units, specifically:

- at least one measurement unit for measuring the thickness distribution or X-Y position of the thickness peaks and
- an application apparatus for controlled change, especially equalizing of convex bows of the substrate stack or of thickness peaks of the substrate stack, especially of the interconnect layer.

Measuring Unit

The first unit is a measuring unit for measuring the local thicknesses along one surface of a substrate stack, especially along one surface of the product substrate. The in particular total thickness variation and/or X-Y position of thickness peaks and/or thickness valleys can be determined from the measurement of the thickness at different X-Y positions along one of the aforementioned surfaces. The X-Y positions are preferably filed in a central control unit for controlling all method steps and components of the device.

The measuring unit consists in particular of a transmitter and a receiver. Preferably the transmitter and receiver are mechanically coupled to one another. In one still more preferred embodiment the transmitter and the receiver are installed jointly in one unit, in particular in a preferably closed housing. Hereinafter, in simplified terms a transmitter is addressed. It means a component which emits signals, for example electromagnetic waves or acoustic sound waves, receives, evaluates and analyzes them in order to determine the aforementioned thicknesses as values which can be further used for the equalizing as per embodiments of the invention.

In a first embodiment of the measuring unit of the invention it is an optical measuring unit, in particular with an optical transmitter. The optical measuring unit preferably has an interferometer. The interferometer can be a monochromatic, or according to one preferred embodiment, a polychromatic white light interferometer. In one special embodiment the interferometer is an infrared interferometer.

In a second embodiment of the invention the measuring unit is made as an acoustic measuring unit, in particular with an acoustic sensor. The acoustic measuring unit preferably has one ultrasonic transmitter. To inject the ultrasonic signal into the

average roughness, the quadratic roughness or for the averaged peak-to-valley height.

The roughness of the mounting base as per embodiments of the invention is as small as possible, in particular less than 100 μm , preferably less than 10 μm , more preferably less than 1 μm , most preferably less than 100 nm, most preferably of all less than 10 nm, still more preferably less than 1 nm. The mounting base consists preferably of a mechanically rigid material with low plastic deformability, especially granite. The modulus of elasticity of the mounting base is especially greater than 0.01 GPa, preferably greater than 0.1 GPa, still more preferably greater than 1 GPa, most preferably greater than 100 GPa, most preferably of all greater than 1000 GPa. The high surface evenness and thus low roughness are achieved in particular by grinding, polishing and etching methods.

The wafer chuck moves preferably on air-cushion bearings along the mounting base. The surface of the wafer chuck preferably has a correspondingly low roughness. The roughness of the wafer chuck surface is in particular less than 100 μm , preferably less than 10 μm , more preferably less than 1 μm , most preferably less than 100 nm, most preferably of all less than 10 nm, still more preferably less than 1 nm. The modulus of elasticity of the wafer chuck is especially greater than 0.01 GPa, preferably greater than 0.1 GPa, still more preferably greater than 1 GPa, most preferably greater than 100 GPa, most preferably of all greater than 1000 GPa.

In one especially preferred embodiment the position-dependent thicknesses are determined in a two-step process. In the first step the thicknesses of the substrate stack are determined by relative movement between the substrate stack and the transmitter, including their X-Y positions. Afterwards a 180° rotation of the substrate stack around its surface normal and repeated measurement of the substrate stack take place. The

actual thickness distribution of the substrate stack can then be computed from the arithmetic mean of the two thickness distributions before and after the rotation. The prerequisite for this averaging process is the unambiguous assignment of the substrate stack points and X-Y positions between the two coordinate systems which are turned to one another. The advantage of this embodiment is mainly that unevenness of the wafer chuck and/or of the mounting plate has less effect on the averaged end result by statistical averaging of the measured layer thicknesses. If repeated measurement of each point were carried out without the rotation as per embodiments of the invention, only better counting statistics would be obtained.

Application apparatus

The second unit or device which is in particular independent as per embodiments of the invention is an application apparatus by which the substrate stack is exposed to local influences, in particular deformations, to the extent one is necessary for reducing the TTV of the substrate stack. The influences can be mechanical, electrical, magnetic and/or thermal influences. As per embodiments of the invention mechanical and/or thermal influences are preferred.

The thickness distribution of the substrate stack which was prepared beforehand and which was filed in the especially software-controlled control apparatus is now used to recognize local extremes (therefore peaks and/or valleys). These extremes occur in particular in substrate stacks which have been produced by a temporary bond. At least predominantly the substrate stack points to locally different thicknesses due to nonuniformity of the thickness of the temporary cement or of the interconnect layer. Furthermore the thickness, especially in isolation or subordinated, is

influenced by the topography of the product substrate, by inclusions, particles or thermal loading.

One embodiment of the invention consists in using the application apparatus to reduce local extremes in a controlled manner. In particular the application apparatus is designed to be rigid analogously to the measuring unit, is manufactured accordingly exactly, can be loaded, and delivers reproducible results. In special embodiments the application apparatus is combined with a measuring unit, especially located and fixed on the same mounting base. Thus, a compact embodiment of the system is devised. Furthermore the transport path of the substrate stack between the measuring unit and the application apparatus is minimal. Another advantage of the combination of the measuring unit and the application apparatus is the parallel execution of the measurement and local change of the thicknesses.

It would in particular be conceivable for the transmitter of the measuring unit to be located at a first position and an application unit of the application apparatus to be located at a second position of the embodiment of the invention. To the extent the substrate stack is moved over the mounting base such that the X-Y position of a local peak of the substrate stack first passes the position of the transmitter of the measuring unit and afterwards the position of the application unit, at a first position a thickness measurement can be taken and at a second position simultaneously the change of the layer thickness can be carried out. Thus a correspondingly higher throughput is achieved.

In an entirely special embodiment the measuring unit and the application apparatus are located separately from one another, preferably in different modules of the same plant (Eng.: fab) or even in modules of different fabs. In this way the measurement of the

thickness distribution is separated from the equalizing of the substrate stack and as per embodiments of the invention it becomes possible to combine measuring units and application apparatus of different manufacturers with one another.

The application apparatus can be made as a spot application apparatus or as a surface application apparatus. The surface application apparatus is the embodiment which is preferred since the throughput is distinctly higher for the equalizing. It consists in particular of a field of spot application apparatus, and each of the spot application apparatus can be in particular triggered and monitored separately. The manner of operation of the method as per embodiments of the invention and of the device as per embodiments of the invention becomes in particular clear and simple by the description of the manner of operation of a spot application apparatus.

Spot Application apparatus

In a first embodiment of the invention the application apparatus has exactly one application unit which is able to apply in particular mechanical pressure to the surface of the substrate stack, especially to the surface of the product substrate. The application unit is preferably stamped to be pin-shaped. Preferably the compressive stress on the convex bow of a local peak causes flow of the interconnect layer between the carrier substrate and the product substrate and thus a level of the thickness of the interconnect layer and thus of the substrate stack. The tip of the pin can have different shapes. Thus, pointed shapes, rounded, especially spherical, or rectangular shapes are conceivable.

In one embodiment of the invention the application unit can be heated. The temperature of the application unit can be regulated in particular with a precision better than 5°C, preferably better than 2°C, still more preferably better than 1°C, most preferably better than 0.1°C, most preferably of all better than 0.01°C. The triggerable temperature range of the application unit is especially between room temperature and 100°C, preferably between room temperature and 300°C, still more preferably between room temperature and 500°C, most preferably between room temperature and 700°C. The heatable application unit can either make contact with the surface of the substrate stack and deliver heat via heat conduction, or can be stopped at a certain distance from the surface of the substrate stack and can heat the more immediate vicinity by heat radiation and/or thermal convection. The delivery of heat leads mainly to a (locally limited) heating and thus a change of the viscosity, especially a decrease of viscosity, of the temporary cement and thus preferably to a decrease of convex bows by relaxation, especially due to the internal stress of the product substrate and/or carrier substrate.

In another embodiment of the invention, the application unit has a laser which can be in particular focused as exactly as possible. The coherent laser light is concentrated by optics preferably at a focus which is located between the product substrate and the carrier substrate, therefore in the interconnect layer. In this way heating of the interconnect layer as locally limited, intense and rapid as possible is effected. The laser is preferably pulsed, in particular by continuous, preferably high- frequency turning on and off. This pulsing for the most part completely prevents heating of the vicinity. The laser is in particular an infrared laser, a laser for visible light or a laser for UV light. The wavelength of the laser is preferably between 10^{-8} and 10^{-3} m, preferably between 10^{-8} and 10^{-4} m, most preferably between 10^{-8}

and 10^{-5} m. Lasers in the wavelength ranges which can cause a chemical reaction in the interconnect layer are especially preferable. This chemical reaction leads in particular to breakdown of a covalent bond. The energy which is necessary for breakdown of a chemical bond is preferably in the UV wavelength range. This chemical reaction can possibly influence the viscosity behavior of the interconnect layer and thus can further optimize the reduction of the TTV.

In one especially preferred embodiment, the application unit combines a heating unit and/or pressure unit and/or laser unit so that thermal and/or mechanical and/or photochemical stresses are possible at the same time.

Other embodiments of the invention which however are not preferred call for switching the application unit as a field electrode so that an electrical field of high intensity can be generated at a tip which is as pointed as possible. In doing so the wafer chuck is set in particular to a potential which is accordingly charged oppositely. The interconnect layer can be physically and/or chemically influenced by the generation of correspondingly high electrical fields.

In another embodiment of the invention, it is conceivable to build the application unit as a magnetic pole shoe so that a magnetic field as strong as possible can be generated. This field penetrates the substrate stack and causes changes in the interconnect layer which allow equalizing.

As per embodiments of the invention the spot application apparatus and/or the wafer chuck are controlled via corresponding software and/or firmware and/or hardware such that the convex bulges or thickness peaks which are to be applied upon are approached individually

and are planarized as extensively as possible by the loading of the application unit on the surface. This minimizes the TTV of the substrate stack.

The application unit according to one preferred embodiment of the invention in any embodiment can be moved at least along the Z axis.

Surface application apparatus

Surface application apparatus consist of several application units which can be triggered in particular separately and/or which are arranged discretely. Each individual application unit, as described above, can be operated therefore mechanically and/or thermally and/or electrically and/or magnetically in order to influence the convex bows or thickness peaks of the substrate stack and thus to minimize the TTV. In contrast to the spot application apparatus, surface application apparatus can control several application units at the same time and thus can treat several locations, in particular several convex bows, at the same time. This leads to a considerable acceleration, especially by a factor corresponding to the number of application units provided.

The following embodiments of the invention are devices which consist of several application units. The application units are arranged next to one another, in particular flush in a rest position. The arrangement of the application units is preferably regular, in particular the arrangement takes place along a rectangular grid. An arrangement of the application units in several circles concentric to one another with a radius which increases accordingly would also be conceivable. Another

possibility which is conceivable as per embodiments of the invention would be an arrangement of the application units along a hexagonal or surface-centered grid.

According to a first embodiment of the surface application apparatus of the invention, the application units apply directly on the surface of the product substrate. The application units can mechanically and/or thermally and/or electrically and/or magnetically change the surface of the product substrate. Analogous considerations as in the above described spot application apparatus apply to the different versions of influence. Accordingly the application units can in turn have built-in heaters and/or apparatus for producing strong electrical and/or magnetic fields. The use of static application units or those which do not apply a mechanical load and which do not change their position or in any case change it to adjust a distance to the surface of the product substrate, therefore always have a distance greater than zero to the surface of the substrate stack, would also be conceivable. The convex bows in this embodiment are influenced largely, in particular exclusively, by heating elements. This surface application apparatus would be limited in particular to locally adjustable heaters which compared to the above described embodiment would not lead to optimum results.

In another embodiment of the invention a membrane which is flexible and stretchable in the application direction of the application units is stretched in front of the application units. The membrane is at least locally deformed by the application units into one application position. The membrane, in contrast to the discrete application units, has a continuous surface, especially one which covers one entire application side of the surface application apparatus.

This embodiment leads to an improved planely loading of the surface of the product substrate/substrate stack.

The membrane can fundamentally consist of any material, but is preferably wear-resistant, temperature-resistant, elastic and/or has correspondingly high strength.

For hardness there are different technological characteristics. The correct hardness measurement method depends on many influencing factors. The most important factors are the material which is to be tested as well as the test body. Metals and ceramics, therefore solids with correspondingly high strength and/or the corresponding capacity for plasticity, are primarily but not exclusively tested with Rockwell, Brinell and Vickers hardness methods. The conversion of the individual hardness measurement values is conditionally possible. Corresponding tables and formats exist and are known to one skilled in the art. But it must be mentioned that exact conversion is not always possible or is inexact. The following hardness measurement values relate to Vickers hardness. The Vickers hardness of the membrane material is in particular greater than 100, preferably greater than 500, still more preferably greater than 1000, most preferably greater than 5000, most preferably of all greater than 10000. In particular embodiments of the invention, membrane materials can also be preferably used which have a very low hardness and therefore can be especially well adapted to the surface of the substrate. These especially preferred membrane materials consist mainly of plastics, especially of plastics which are temperature-stable up to several hundred degrees Celsius. The preferred hardness measurement method for plastic is the Shore method. The Shore hardness of plastics is in particular less than 100, preferably less than 75, still more preferably less than 50, most preferably less than 25, most preferably of all less than 5. The

exact definition can be looked up in the corresponding standards DIN ISO 7619-1 and DIN EN ISO 868.

In this connection temperature resistance means that the membrane at elevated temperature does not begin to creep, therefore that no plastic deformation by diffusion processes arises under load. In order to avoid having to address the more detailed physical and chemical processes of creep, a creep-resistant material as per embodiments of the invention is characterized in that the corresponding material does not experience plastic deformation as a function of time up to a maximum tension and a maximum temperature. The maximum temperature as claimed in the invention is in particular greater than 0°C, preferably greater than 50°C, still more preferably greater than 100°C, most preferably greater than 250°C, most preferably of all greater than 500°C. The maximum tension as per embodiments of the invention is in particular greater than 10E6 Pa, preferably greater than 10E7 Pa, more preferably greater than 10E8 Pa, most preferably greater than 10E9 Pa.

The elasticity is described by the modulus of elasticity. The modulus of elasticity of the membrane as per embodiments of the invention is especially between 1 GPa and 1000 GPa, preferably between 10 GPa and 1000 GPa, more preferably between 25 GPa and 1000 GPa, most preferably between 50 GPa and 1000 GPa, most preferably of all between 100 GPa and 1000 GPa. The modulus of elasticity of some types of steel is for example roughly 200 GPa.

The yield point specifies the maximum tension at which the body does not yet deform plastically. As per embodiments of the invention, materials for the membrane are preferred which have a high yield point in order to prevent permanent plastic deformation. The yield point of the membrane material as per embodiments of the invention is especially greater than 1 MPa, preferably greater than 10 MPa, still more preferably greater than 100 MPa, most preferably greater than 500 MPa, most preferably of all greater than 1000 MPa.

Membrane materials are especially preferred which do not contaminate the surface of the product substrate. The membrane materials should therefore preferably be free of all elements which are critical to semiconductor applications, especially metals like copper, gold, molybdenum and titanium. A membrane material is preferably used which is based on a polymer and/or carbon material.

In another embodiment of the invention the application units are pressure chambers which are partitioned off from one another and which can be triggered and/or supplied especially separately with pressure. The pressure chambers are surrounded in the direction of the surface of the product substrate which is to be applied upon by a membrane and laterally by elastically deformable pressure chamber walls. The pressure chambers can preferably be exposed to overpressure and/or underpressure individually and allow deformation of the membrane and thus upon contact with the surface of the product substrate allow the application on the convex bows as per embodiments of the invention.

The pressure chambers as per embodiments of the invention have in particular a pressure between 10^{-8} mbar and 10 bar, preferably between 10^{-8} mbar and 8 bar, still more preferably between 10^{-8} mbar and 6 bar, most preferably between 10^{-8} mbar and 4 bar, most preferably of all between 10^{-8} mbar and 2 bar. The pressure chambers can be supplied with any fluid, therefore with liquids and/or gases. Accordingly they are hydraulic and/or pneumatic pressure chambers. Preferably liquids are not used, so that pneumatic pressure chambers are preferred.

The combination of a surface application apparatus with a measuring unit is conceivable as per embodiments of the invention, as described above. Preferably first the entire substrate stack (in particular

with respect to its thickness and/or the thickness of individual layers of the substrate stack such as the thickness of the product substrate, carrier substrate and/or interconnect layer) is measured. Afterwards the improvement of the thickness variation (TTV) as per embodiments of the invention takes place in the surface application apparatus as per embodiments of the invention. In another optional step, the repeated measurement of the substrate stack then takes place. If the determined TTV should be smaller than a given boundary value, the substrate stack can be further processed. If it is greater, a repeated improvement of the thickness variation (TTV) as per embodiments of the invention takes place by means of the application apparatus as per embodiments of the invention.

In one preferred embodiment the surface application apparatus of the invention and the measuring unit are located within one unit or within the same module, as a result of which the change between the measurement of the substrate stack and the improvement of the TTV is accelerated.

Arrangement and Shape of the Application units

The shape of the application units, in particular on an application side and/or a jacket surface which is located on the end, can be fundamentally optional, to the extent in this way application on the surface is possible. It is shown however that certain shapes are suitable for special applications.

The jacket surfaces of the application units are preferably shaped as round, especially circular, or as square cylinders.

The application side (or application head) is shaped in particular to be rounded, to be flat, pointed, to be conically flattened, or to be hexagonally flat.

Flat application heads are better suited as per embodiments of the invention for reducing convex bows and for heat transfer.

Pointed application heads are mainly important when high electrical and/or magnetic fields are to be produced and localized.

Flat application heads with a hexagonal shape are in particular preferred since they are on the one hand flat, can completely cover one surface and due to their hexagonal symmetry have an optimum radial relation to their corresponding neighbors.

Other advantages, features and details of the invention will become apparent from the following description of preferred exemplary embodiments and using the drawings.

Figure 1a shows a schematic cross section of an ideal substrate stack without TTV,

Figure 1b shows a schematic cross section of a real substrate stack with thickness variation (TTV),

Figure 1c shows a schematic cross section of a treatment process (here: back-thinning) of a real substrate stack,

Figure 2 shows a schematic cross section of one embodiment of a measuring unit as claimed in the invention,

Figure 3 shows a schematic cross section of one embodiment of a spot application apparatus as claimed in the invention,