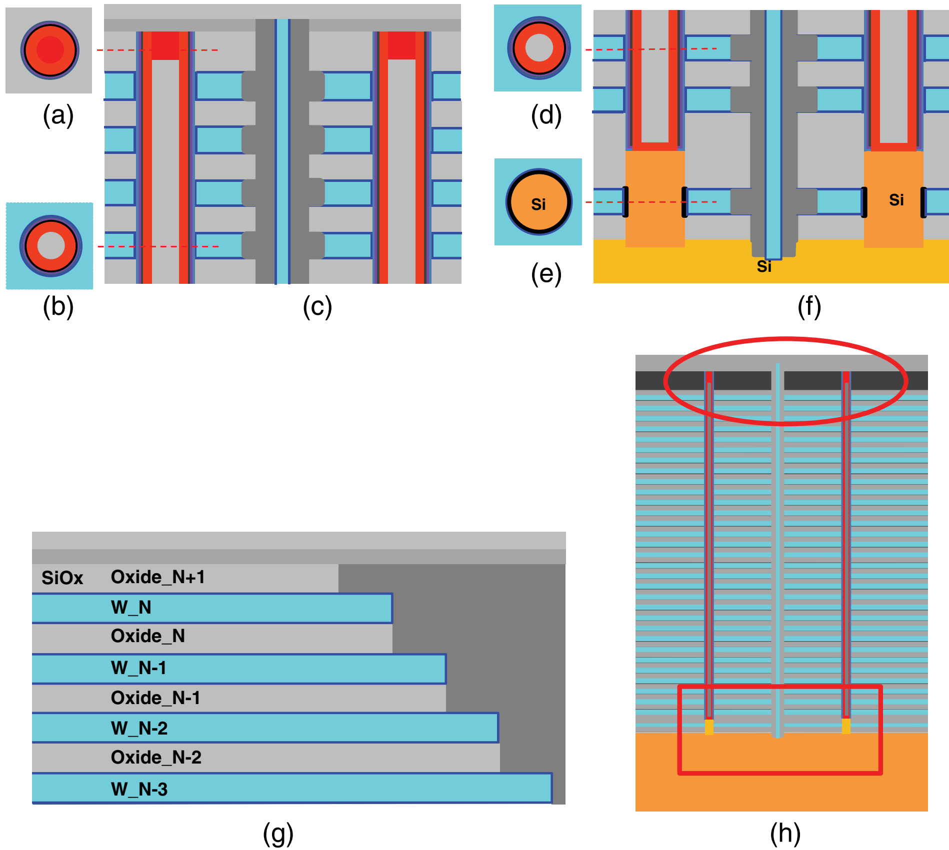



**Figure 2.33** Top-down view (a) in cap oxide and (b) in nitride_N-2; (c) cross-section near the top of the channel; top-down view (d) in nitride_3 and (e) nitride_1; (f) cross-section along the channel at the bottom; (g) cross-section at the top of the staircase; and (h) cross-section of the whole stack.

## 2.2.5 Contact and interconnection modules

At this point, the front-end processes are finished, and all devices in both the peripheral and array areas are built. This section discusses the BEOL processes—contact and interconnect—of 3D-NAND flash. First is the contact module, in which contact holes are etched in the staircase area between the cell and periphery; they land on the W staircases and silicon substrate, creating the connection between the word lines in different layers and the source line in substrate.

After hard mask deposition and photoresist coating, the contact mask is applied [Fig. 2.34(a)]. At first the hard mask is etched and then used to etch contact holes in oxide with etch chemistries that have high selectivity to tungsten [so that the etch process stops whenever the contact holes reaches the tungsten surface in the shallower holes while continuing in the deeper holes, as

**Table 2.4** Process steps for the isolation module of 3D-NAND.

| | |
|---|---|
| Wafer clean | Trench W removal |
| Isolation mask (Fig. 2.26) | Trench TiN removal (Fig. 2.32) |
| Etch hard mask | Wafer clean |
| Etch trenches in ONON multi-layers and stop on silicon | Oxide deposition |
| Remove hard mask [Figs. 2.27(c) and 2.28] | Oxide etch back (Fig. 2.32) |
| Remove nitride layers (Fig. 2.29) | TiN deposition |
| Wafer clean | W deposition |
| Oxidation of SEG (Fig. 2.30) | W CMP |
| TiN deposition | Oxide cap deposition (Fig. 2.33) |
| W deposition (Fig. 2.31) | |

shown in Fig. 2.34(b)]. Chapter 9 of Xiao[11] discusses this etch chemistry used for contact etches with different hole depths.

Because the depth of the contact holes vary between different layers, it would be very difficult to etch all of them with almost 40 different depths in one etch process; multiple masks would be necessary. Depending on the process, ~10 different contact-hole depths can usually be etched in one etch process, and thus four masks and four etch processes are needed to etch all of the contact holes in the staircase and periphery of a 32-cell-stack 3D-NAND flash device. Figure 2.34(c) shows the cross-section after staircase contact etch and hard-mask strip and clean.

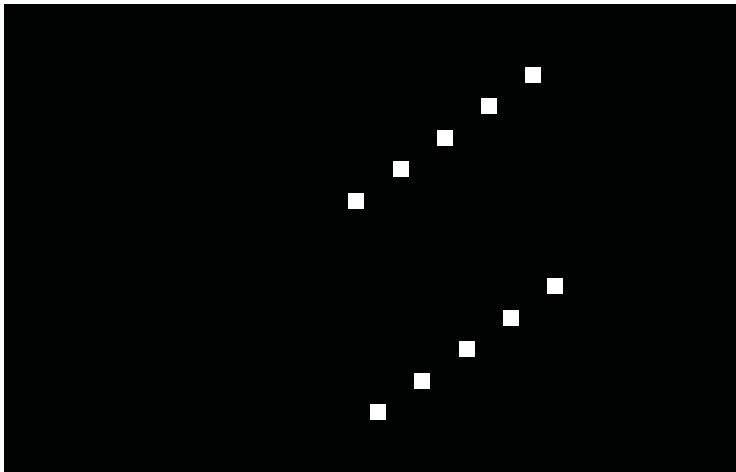
When all of the contact holes have been etched, the wafer is cleaned to remove the polymer residue at the bottom of the contact holes. After sputtering etch removes the native oxide, barrier TiN is deposited [Fig. 2.35(a)], followed by W deposition [Fig. 2.35(b)]. A WCMP process removes W and TiN from the surface [Fig. 2.35(c)], which completes the contact module.

Figures 2.36(a)–(c) illustrate the cross-section of TiN deposition, W deposition and WCMP in cell, staircase and periphery areas, respectively. Figures 2.35(a)–(c) are close-up versions of Figs. 2.36(a)–(c) at the top of the staircases, respectively.

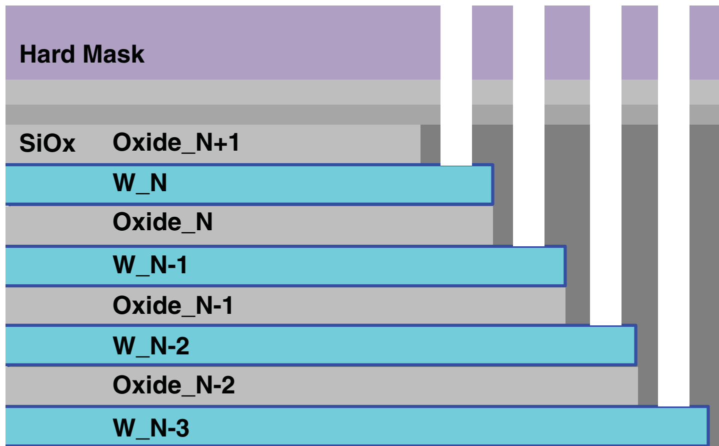
The next module is metal 1, which is a dual-damascene process that forms the local interconnect. It includes two masks, via 1 (V1) and metal 1 (M1). A layer of oxide is deposited to cap the contact W plugs. In the first via process, a V1 mask, which is like a channel mask plus a contact mask (Fig. 2.37), is applied first. Via holes are then etched to land on channel polysilicon plugs and contact tungsten plugs (Fig. 2.38).

Metal 1 (M1) forms the local interconnect in the array area and peripheral area. The pattern in the staircase area is almost the same as V1. Figure 2.39(a) shows the M1 mask, and Fig. 2.39(b) shows the M1 overlaps with the channel, isolation, contact, and V1.

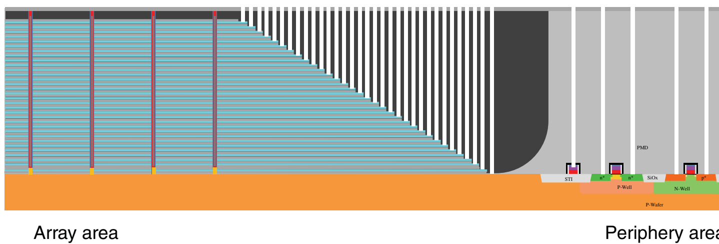
After applying the M1 mask, an oxide etch process is performed to form the trenches of the local interconnect. After photoresist strip and clean, a TiN liner and W are deposited to fill the M1 trenches and V1 holes. WCMP



(a)

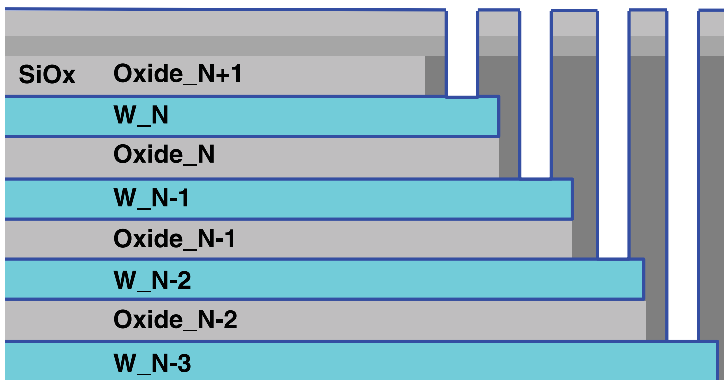



(b)

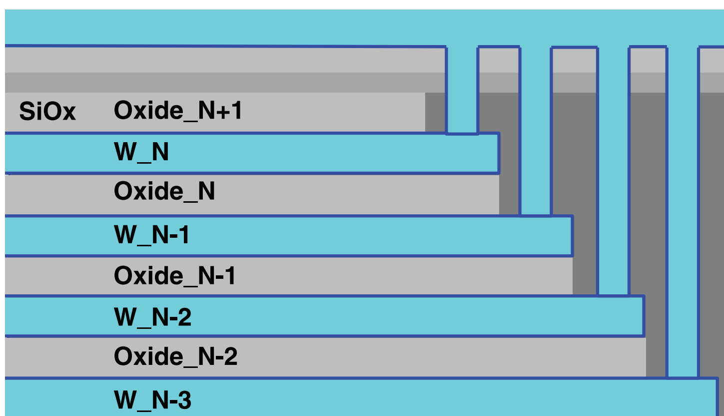



(c)

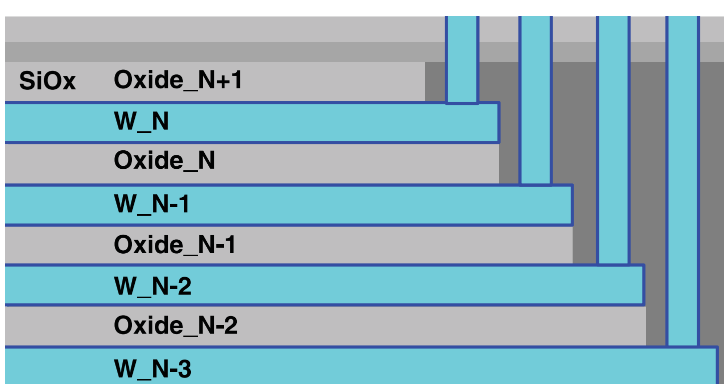
**Figure 2.34** (a) Staircase contact mask, (b) close-up of the top layer, and (c) cross-section of the cell and periphery contacts.




(a)




(b)




(c)

**Figure 2.35** Close-up of the contact processes near the top of the staircase: (a) after TiN deposition, (b) after W deposition, and (c) after WCMP.

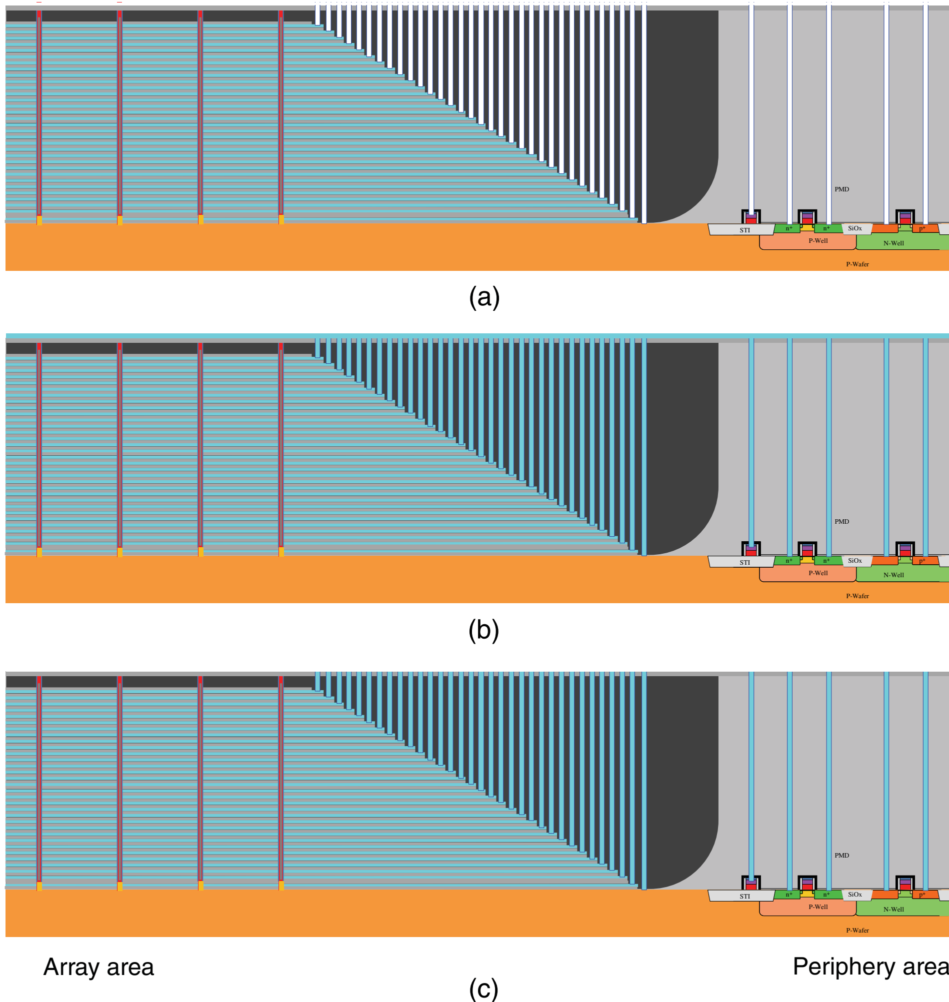



**Figure 2.36** Contact process steps: (a) after TiN deposition, (b) after W deposition, and (c) after WCMP.

removes the W and TiN from the wafer surface and forms the W wires and plugs that connect the channel plugs and contact plugs, as shown in Fig. 2.40.

Metal 2 forms the bit line in the array area, source line and word line wires in the staircase area, and interconnection in the peripheral area. Because each WL between the two isolation walls has four rows of channel holes, the bit lines, which are perpendicular to the WL, must be split into a pitch density four times higher than the channel hole pitches to ensure the registration of a single cell in the string with one bit signal and one word signal. After wafer clean and ILD deposition, there is a via 2 (V2) and metal 2 (M2) process (the via 2 mask is illustrated in Fig. 2.41).

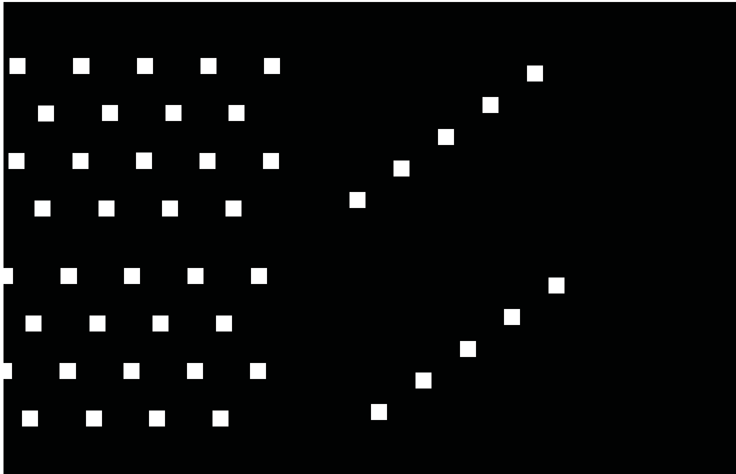


**Figure 2.37** V1 mask.

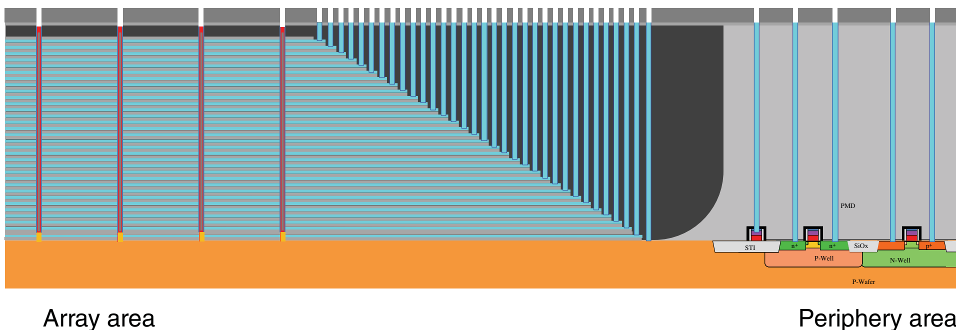



**Figure 2.38** V1 etch in the cell, staircase, and peripheral areas.

After V2 etch, photoresist strip, and clean [as shown in Fig. 2.42(a)], a TiN liner is deposited, followed by W deposition [shown in Fig. 2.42(b)], and then WCMP removes the W and TiN on the surface and forms the V2 W-plugs [Fig. 2.42(c)].

After WCMP, another ILD is deposited, a M2 mask is applied, and metal trenches are etched. Figure 2.43 illustrates the M2 mask, and Fig. 2.44(a) shows the M2 etch. After photoresist strip and clean, a TaN barrier layer and Cu seed layer are deposited into the M2 trenches; after copper plating and anneal, a metal CMP process removes the Cu and TaN from the wafer surface and forms the BL in the array area and WL and SL wires in the staircase area [Fig. 2.44(b)].

Metal 3 (M3) is the last metal layer; it forms interconnect and bond pads. It is usually formed by a stack of metals: a Ti barrier layer at the bottom, an Al-Cu alloy bulk layer, and TiN ARC on top. Via 3 (V3) is a tungsten plug that connects to M2.