

林训涛, 韩国军. 融合沟道倾斜角的3D NAND存储单元建模及编程策略优化[J]. 广东工业大学学报, 2026, 43(2): 91-100. doi: 10.12052/gdutxb.250167.
Lin Xuntao, Han Guojun. Modeling of 3D NAND memory cells with taper angle and programming strategy optimization[J]. Journal of Guangdong University of Technology, 2026, 43(2): 91-100. doi: 10.12052/gdutxb.250167.

融合沟道倾斜角的3D NAND存储单元 建模及编程策略优化

林训涛, 韩国军

(广东工业大学 信息工程学院, 广东 广州 510006)

摘要: 随着3D NAND闪存堆叠层数的不断增加, 深孔刻蚀工艺造成的深宽比也相应增加, 由此形成的沟道倾斜角导致存储单元阈值电压分布展宽, 影响了器件存储可靠性。当前针对该问题的解决方案主要包括刻蚀优化、沟道补偿等工艺方法, 以及纠错编码、读重试等策略, 然而这些方法普遍面临实现复杂、硬件开销大等问题。为克服上述局限性并有效地降低倾斜角的影响, 本文首先基于Sentaurus TCAD平台构建了含倾斜角的存储串仿真模型, 验证分析了深孔倾斜角对阈值电压分布梯度的影响, 在此基础上, 提出了一种相邻单元导通电压幅度动态调整的策略, 该策略根据目标存储单元在存储串中的垂直位置, 差异化调整其相邻字线的导通电压值, 以精准补偿因沟道倾斜引起的纵向电压梯度。仿真结果表明, 本文所提出的编程策略与传统编程策略相比, 在线性调节模式下, 最大电压差与分布标准差分别降低了62.4%与61.4%; 在非线性调节模式下, 二者分别降低了63.7%与64.3%, 同时将相邻单元的阈值电压扰动控制在3 mV以内, 且在不同编程时间下写入速率保持稳定。

关键词: 3D NAND闪存; 器件建模; 沟道倾斜角; 阈值电压分布; 动态导通电压

中图分类号: TP333

文献标志码: A

文章编号: 1007-7162(2026)02-0091-10

Modeling of 3D NAND Memory Cells with Taper Angle and Programming Strategy Optimization

Lin Xuntao, Han Guojun

(School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: With the continuous increase in the number of stacking layers in 3D NAND flash memory, the aspect ratio caused by deep hole etching processes also increases accordingly. The resulting taper angle leads to a broadening of the threshold voltage distribution of memory cells, adversely affecting device storage reliability. Current solutions to this problem primarily include process-based methods such as etching optimization and channel compensation, or strategies like error correction coding and read-retry. However, these approaches generally face challenges such as implementation complexity and high hardware overhead. To overcome these limitations and effectively mitigate the impact of the taper angle, a simulation model of a memory string incorporating taper angles is first constructed based on the Sentaurus TCAD platform, validating and analyzing the influence of deep hole taper on the threshold voltage distribution gradient. On this basis, a strategy for dynamic optimization of the pass voltage amplitude in adjacent cells is proposed. This strategy differentially optimizes the pass voltage of word lines adjacent to the target memory cell according to its vertical position in the memory string, thereby accurately compensating for the longitudinal voltage gradient induced by channel taper. Simulation results demonstrate that compared to conventional programming strategy, the proposed strategy optimizes the maximum

收稿日期: 2025-09-18 录用日期: 2025-10-25 网络首发日期: 2025-12-27

基金项目: NSFC-广东省联合基金资助项目(U2001203)

作者简介: 林训涛(2000-), 男, 硕士研究生, 主要研究方向为3D NAND闪存器件建模与NAND特性分析, E-mail: 1986254495@qq.com

通信作者: 韩国军(1974-), 男, 教授, 博士, 主要研究方向为信息论与差错控制编译码技术、储存器件与系统、智能通信与车联网等, E-mail: gjhan@gdut.edu.cn

voltage difference and the standard deviation of distribution, achieving reductions of 62.4% and 61.4%, respectively, in linear mode, and 63.7% and 64.3% in nonlinear mode, while constraining the threshold voltage disturbance of adjacent cells within 3 mV and maintaining stable programming speed.

Key words: 3D NAND flash memory; device modeling; channel taper angle; threshold voltage distribution; dynamic pass voltage

3D NAND闪存技术通过垂直堆叠存储单元,突破了传统平面闪存的物理缩放极限,已成为高密度固态存储的主流方案^[1]。随着大数据、人工智能与云计算技术的快速发展,对存储密度和性能的要求不断提高,这推动闪存技术持续创新。在多种架构中,基于电荷陷阱型(Charge-Trap, CT)的替代栅或通道孔工艺因集成特性优异,被广泛应用于3D NAND制造^[2-3]。与传统浮栅结构相比,CT结构具有更好的抗串扰能力和更简单的工艺集成性,在高密度存储中优势显著^[4]。

然而,随着堆叠层数增加,深孔刻蚀工艺会引入明显的几何形变,使存储串(Memory String)沟道与垂直方向形成倾斜角。该倾斜角不仅影响器件电学特性,还降低产品良率和可靠性^[5-6]。这一非理想效应破坏存储串的结构对称性,引起串内不同位置单元物理参数的系统性变异。具体表现为沟道半径沿深度方向呈梯度分布:顶部单元沟道半径较大,底部逐渐减小。这种纵向差异直接影响各单元的栅控特性,导致同一存储串内阈值电压(V_{th})出现显著差异^[6]。此外,倾斜角还改变单元间耦合电容分布,加剧电学干扰的复杂性。上述多物理参数变异共同导致阈值电压分布显著展宽^[7-10],成为高层数3D NAND器件性能与可靠性提升的关键瓶颈。

阈值电压分布的均匀性直接决定多级存储的操作容限与数据完整性。在多级存储中,每个单元需存储多个比特数据,要求将有限阈值电压窗口划分为多个狭窄状态。分布展宽会加剧读干扰,导致电压状态重叠,增加读错误率^[11-12];同时也会缩窄有效存储窗口(Memory Window),降低数据区分度,最终减少存储器寿命^[6,13-14]。在高密度存储中,即使微小分布展宽也可能导致比特错误率显著上升,严重影响产品可靠性。

目前对该问题的研究主要围绕两个方向:一是通过工艺优化减小倾斜角或补偿其影响,二是通过电路设计与纠错码进行后期补偿。工艺方面,通过改进刻蚀参数、优化掩模材料、采用新型刻蚀试剂等方法减小倾斜角;或通过修改沉积工艺,在沟道成型时沉积具有锥形轮廓的材料层以补偿电学变异^[6]。这些方法虽有一定效果,但普遍面临高工艺成本与技术

瓶颈。随着堆叠层数继续增加,深宽比不断增大,工艺难度急剧上升,对沉积精度要求极高。电路方面则采用纠错编码(Error Correction Code, ECC)^[15-16]、读重试算法^[17-18]等策略补偿电压分布不均匀。这些方法可在一定程度上提升数据可靠性,但往往会增加芯片面积、功耗和时延,在高密度存储应用中ECC电路的开销可能难以接受。

从器件物理层面深入理解倾斜角的影响机理,并据此提出有效的优化调控方法,具有重要的理论价值与现实意义。通过建立精确的物理模型,可以系统地研究倾斜角对器件电学特性的影响规律,为工艺优化提供理论指导。同时,基于物理机理提出的电路优化方案往往具有更高的效率和更好的适用性,能够在不过多增加硬件开销的情况下显著改善器件性能。

本文基于Sentaurus TCAD仿真平台,构建了10层3D NAND闪存的存储串模型,重点研究了沟道倾斜角对阈值电压分布的影响规律。通过建立包含精确几何形变的仿真模型,并校准关键物理参数,系统分析了不同倾斜角度下串内各单元电学特性的变化。仿真结果验证了沟道倾斜角导致的纵向阈值电压梯度的现象:由于沟道半径自上而下递减,顶部单元阈值电压降低,而底部单元阈值电压升高,造成阈值电压分布展宽。

针对该问题,本文提出一种位置相关的导通电压调节方法,根据目标编程单元在存储串中的位置,动态调整其相邻字线的导通电压值,以针对性补偿沟道倾斜带来的电学差异。仿真表明,该优化方法可显著改善阈值电压分布的均匀性,降低标准偏差与最大电压差。同时,该方法仅通过电气调控实现性能优化,无需改变现有工艺,具有较强的工程实用性与可实施性,为高层数3D NAND存储器的可靠性提升提供了一种有效的电路级解决方案。

1 3D NAND闪存简介

本节将叙述3D NAND闪存的基本结构和器件层面的关键操作,并详细阐述沟道倾斜角的影响机制。

1.1 3D NAND闪存基本结构

3D NAND闪存的存储阵列采用独特的立体堆

叠结构以实现高密度数据存储。如图1所示,其基本组成单位为闪存块(Block),每个块包含大量存储单元(Cell)。

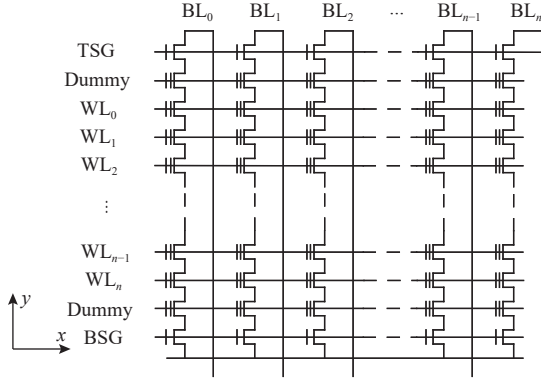


图1 闪存阵列基本结构

Fig.1 Fundamental architecture of flash memory array

在结构组织上,沿 x 方向排列且共享同一控制栅的多个存储单元通过字线(Word Line, WL)互联,字线用于控制同一行所有单元的读写擦操作;沿 y 方向排列且串联于同一位线上的存储单元组成一个串。在每个存储串的首尾两端设置有虚拟单元(Dummy Cell),用于抑制边缘存储单元的热载流子注入效应,改善电场分布均匀性。每个串连接1条位线和1条源线,构成独立的电流通路。多个串与字线在三维空间中交叉形成块,它是执行擦除操作的基本单位。最终,若干个块共同组成1个逻辑单元号,作为设备内部独立管理和寻址的基本功能模块^[19]。

1.2 编程操作与擦除操作

1.2.1 编程操作

在3D NAND闪存的编程操作期间(传统编程方案),各端口的电压配置如图2所示。对选定的目标编程单元,其对应的字线(WL₁)施加编程电压(V_{pro} ,通常为15~20 V的高压),此电压足以使电子穿过隧穿氧化层注入电荷存储层。与此同时,所有未选中的字线则施加导通电压(V_{pass} ,通常为8~10 V),该电压可使这些单元处于导通状态,确保形成连续的电流通路,但又不足以引起非目标单元的意外编程。

在位线侧,目标编程单元对应的位线(BL₁)被施加0 V电压,形成必要的电势差以促进编程操作;而其他未选中的位线则施加电源电压(V_{CC} ,通常为2~3 V),此电压可关闭这些位线对应的串,防止串扰。选择管方面,顶部选择管(Top Select Gate, TSG)施加 V_{CC} ,使其处于导通状态;而底部选择管(Bottom Select Gate, BSG)则保持接地,共同确保编程操作仅在选定的目标单元上发生,从而实现精确的数据写入。

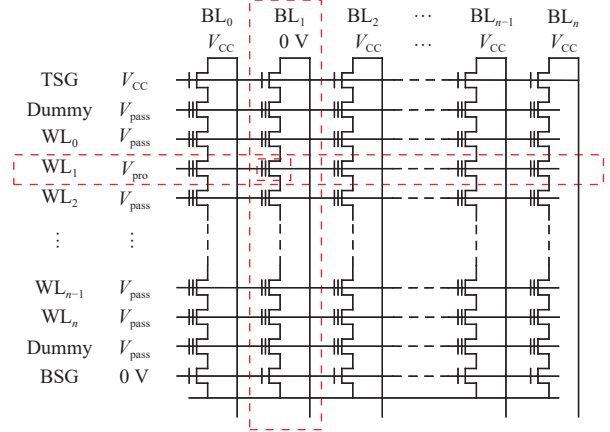


图2 闪存阵列各端口电压(传统编程方案)

Fig.2 Terminal voltages in a flash memory array (conventional programming scheme)

1.2.2 擦除操作

在3D NAND闪存的擦除操作中,为了最大限度地减少外围电路在芯片面积中的占比并提升擦除的效率,目前业界普遍采用体擦除的统一操作方式。该擦除方案的具体操作为:将所有需要擦除的字线统一施加0 V固定电位,同时在器件的硅基底施加1个高压正偏置脉冲(典型幅值为15~20 V)。

在此偏置条件下,沟道区域与基底之间形成极强的垂直电场,该电场使得存储在电荷陷阱层中的电子通过FN隧穿(Fowler-Nordheim Tunneling)机制被强制拉出至衬底,与此同时空穴在电场作用下被注入到电荷存储层中,从而有效中和原有存储电荷,最终实现存储单元的快速擦除。这种基于体效应的擦除机制,无需对每个存储单元进行复杂的单独寻址,可并行对整个存储块内所有单元执行擦除动作,极大提高了擦除操作的效率,是3D NAND架构能够在有限芯片面积内实现高集成度和高可靠性的关键特性之一。

1.3 沟道倾斜角

理想圆柱形沟道的存储串截面如图3(a)所示,其沟道半径保持均匀且一致。图3(b)则显示了引入沟道倾斜角后的实际结构,其中参数 ΔT_{ch} 表征了沟道顶部与底部之间的半径差值。沟道倾斜角的存在对存储串的电学特性产生了重要影响,其作用机制主要体现在以下几个方面。

沟道倾斜角导致存储串的沟道半径沿垂直方向呈现梯度变化。顶部沟道半径较大,而底部沟道半径较小。这种几何形变直接影响了存储单元的导电特性:顶部较大的沟道半径显著降低了沟道电阻,进而增强了载流子的迁移能力,使得沟道更易在较低的

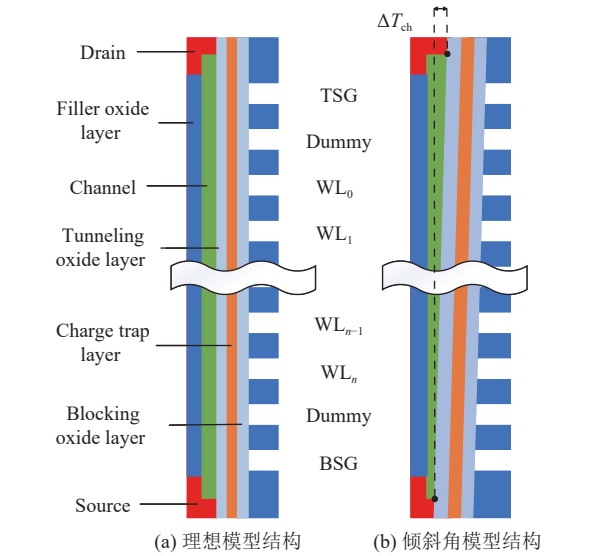


图 3 沟道倾斜角对存储串模型截面结构的影响

Fig.3 Impact of channel tilt angle on cross-sectional structure of memory string model

栅极电压下形成导通;反之,底部较小的沟道半径则增大了沟道电阻,削弱了沟道的导电能力,使得底部存储单元需要更高的栅极电压才能实现完全开启。

同时,倾斜角还破坏了结构对称性,导致单元间耦合电容分布改变,产生额外的编程干扰和读操作的单元间串扰。

这些物理参数变异最终导致存储串的阈值电压沿布线方向出现系统性漂移:顶部单元的阈值电压较低,底部单元的阈值电压较高。这种纵向梯度造成阈值电压分布展宽,严重影响多级存储操作的可靠性。

2 模型搭建与验证

为探究沟道倾斜角对3D NAND存储串阈值电压分布的影响机制,本文基于Sentaurus TCAD平台,建立了精确的存储串仿真模型,系统开展了结构搭建与电学特性的仿真工作,并利用所构建的模型对沟道倾斜角引发的锥角效应及其电学影响进行了验证。

2.1 器件结构与仿真设置

本文建立的3D NAND存储串TCAD仿真模型结构如图4所示。随着3D NAND堆叠层数持续增加,三维仿真所需计算资源急剧增长,为提升计算效率,本文仅对对称截面进行二维建模。该方法在保证关键物理过程可模拟的前提下,显著降低了计算复杂度,是当前高层数3D NAND存储器件仿真中的主流方案^[20]。需要指出的是,此类方法虽高效,但仍无法完全替代真实三维仿真。

该模型基于具有圆柱形沟道的位成本可缩放

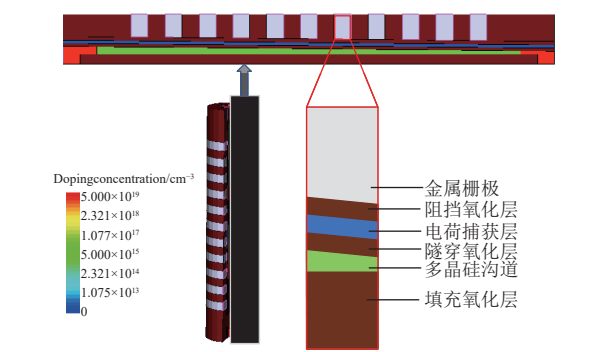


图 4 CT型3D NAND存储串模型三维结构与二维结构

Fig.4 Three-dimensional and two-dimensional structures of charge-trap 3D NAND memory string models

(BiCS)结构^[21-23],包含2个顶部选择管、2个底部选择管、4个虚拟单元以及 n 个存储单元($WL_0 \sim WL_{n-1}$)。虚拟单元设置在存储串的首尾两端。结构中的关键组成部分包括:多晶硅沟道(Polysilicon Channel)、源极和漏极(Source/Drain)、隧穿氧化层(Tunneling Oxide Layer)、氮化硅电荷俘获层(Charge Trap Layer)、阻挡氧化层(Blocking Oxide Layer)以及金属栅极。此外,模型还包含了中心填充层(Filling Oxide Layer)以保持结构完整性。

表1详细列出了器件模型的各项关键结构参数和掺杂配置。横向栅极长度(L_g)和纵向栅极长度(T_g)分别为28 nm和40 nm。电荷俘获层采用经典的阻挡氧化层-氮化硅层-隧穿氧化层结构,厚度分别为6 nm、5 nm和5 nm。漏极侧的多晶硅沟道厚度(T_{ch}^D)为10 nm,源极侧的多晶硅沟道厚度(T_{ch}^S)为5 nm。掺杂配置方面,源极和漏极区域采用高浓度N型掺杂(N_S/N_D),沟道区域采用轻掺杂(N_{ch})。氮化硅电荷存储层中的电荷陷阱(n_{CTN})和空穴陷阱(p_{CTN})浓度均设置为 $5 \times 10^{19} \text{ cm}^{-3}$,以准确模拟电荷俘获行为。

表 1 仿真器件结构参数和掺杂配置

Table 1 Structural parameters and doping profile of the simulated device

参数	符号	值
栅极长度	L_g/nm	28
栅极厚度	T_g/nm	40
阻挡氧化层厚度	T_{BO}/nm	6
氮化硅层厚度	T_N/nm	5
隧穿氧化层厚度	T_{TO}/nm	5
多晶硅沟道厚度(漏极侧)	T_{ch}^D/nm	10
多晶硅沟道厚度(源极侧)	T_{ch}^S/nm	5
漏/源极掺杂浓度	$N_S/N_D/\text{cm}^{-3}$	5×10^{19}
多晶硅沟道掺杂浓度	N_{ch}/cm^{-3}	1×10^{15}
存储层电子缺陷浓度	n_{CTN}/cm^{-3}	5×10^{19}
存储层空穴缺陷浓度	p_{CTN}/cm^{-3}	5×10^{19}

2.2 器件电学特性仿真及配置

在使用Sentaurus结构编辑器(Sentaurus Structure Editor, SDE)完成器件结构搭建和网格划分后,本文采用S-Device(Sentaurus Device)模块对3D NAND存储串进行电学特性仿真分析。针对编程和擦除操作,仿真采用非局域隧穿模型(Non-local Model),该模型在沟道与隧穿氧化层界面处进行了精细网格定义。考虑到电荷俘获层、隧穿层和沟道层的总厚度,将隧穿距离设置为10 nm。

在氮化硅电荷俘获层中,同时考虑了载流子的俘获-发射模型和漂移-扩散模型。具体参数设置为:施主型和受主型缺陷密度均为 $5 \times 10^{19} \text{ cm}^{-3}$,施主陷阱能级位于价带上方1.5 eV,受主陷阱能级位于导带下方1.24 eV^[23]。通过缺陷模型与非局域隧穿模型的耦合,准确模拟了电子隧穿进入氮化硅层后的俘获和脱陷动态过程。

为模拟高电场条件下存储层中陷阱势垒降低导致的电子发射现象,在缺陷模型中引入了Poole-Frenkel模型,该模型能够有效预测高电场下电子从陷阱中发射的概率^[24-25]。对于擦除操作,采用基于栅致漏电(Gate-Induced Drain Leakage, GIDL)机制的擦除方式^[26],通过在沟道区域配置带带隧穿(Band-To-Band Tunneling, BTBT)物理模型,准确模拟了当选择管的源、漏电极与栅极之间存在较大电势差时引发的带带隧穿效应。这一效应会产生电子-空穴对,其中空穴会注入电荷存储层实现擦除过程。

仿真同时采用了掺杂浓度依赖的迁移率模型、肖克利-里德-霍尔(Shockley-Read-Hall, SRH)复合模型、高场饱和模型和迁移率表面退化模型。由于本文主要关注沟道倾斜角对阈值电压分布的影响,而多晶硅沟道中随机晶界效应对研究目标的影响相对较小,因此在电学特性仿真中未考虑多晶硅的随机晶界效应。

2.3 验证锥角效应

为验证本文所建仿真模型对实际3D NAND制造中沟道倾斜效应的表征能力,本文对基于Sentaurus TCAD平台构建的存储串模型进行了系统性仿真测试,通过量化分析锥角效应与阈值电压分布的关联特性,证明模型可准确反映实际工艺中的电学特性变异。阈值电压采用恒定电流法测量,即定义为漏极电流(I_d)为 $1 \times 10^{-6} \text{ A}/\mu\text{m}$ 时对应的栅极电压值(V_g)^[6]。

仿真中的编程操作采用传统编程方案中应用的偏置条件:目标字线施加19 V编程电压,相邻字线施加8 V导通电压,顶部选择管加2 V电压,底部选择管接地,选中位线接地,未选位线施加2 V电压。测试采

用单个单元编程策略,有效避免多单元并行编程引入的串扰和电荷耦合效应。每次测试前执行全串擦除以消除残余电荷,确保阈值电压测量不受历史状态影响,保证数据的准确性和可重复性。

仿真结果充分证明了该模型能够表征不同倾斜角下的锥角效应。图5展示了编程状态下4种倾斜角模型的阈值电压分布对比:理想垂直沟道模型表现出高度均匀的分布特性,各层阈值电压维持在2.627~2.628 V的极窄范围内,标准差低至0.000 5 V,表明模型在无倾斜条件下具有优异的基准一致性。

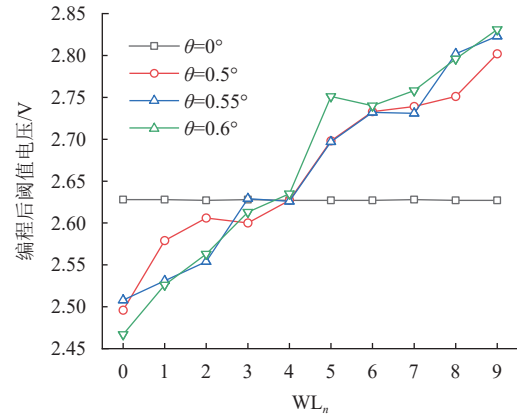


图5 不同沟道倾斜角下的存储串阈值电压分布
Fig.5 Threshold voltage distribution of memory string under different channel tilt angles

随着倾斜角增大,阈值电压分布呈现明显的梯度特征和非线性响应。当 $\theta=0.5^\circ$ 时,阈值电压从顶层 WL_0 的2.496 V递增至底层 WL_9 的2.802 V,阈值电压差值 ΔV_{th} 为0.306 V,标准差 σ_{th} 为0.092 V; $\theta=0.55^\circ$ 时,阈值电压分布范围为2.508~2.823 V, ΔV_{th} 增大至0.315 V, σ_{th} 增至0.108 V; 而 $\theta=0.6^\circ$ 时,阈值电压梯度更加显著,分布范围扩展至2.467~2.831 V, ΔV_{th} 达到0.364 V, σ_{th} 进一步增至0.125 V。

这些数据清晰展示了倾斜角与阈值电压分布偏差之间的正相关关系。特别值得注意的是,随着倾斜角从 0.5° 增加至 0.6° ,阈值电压差值增大了18.9%,标准差增大了35.9%,表明倾斜角增大对阈值电压分布均匀性的影响呈现加速恶化趋势。该结果与理论分析中沟道倾斜角的影响机制高度吻合^[6]。仿真数据呈现的规律性变化表明,所建立的模型能够有效反映因沟道几何形变引起的电学特性变异,准确捕捉到锥角效应导致的阈值电压梯度分布特征。这一验证结果证明,本文采用的仿真方法和模型设置适用于研究3D NAND器件中沟道倾斜角对阈值电压分布的影响机制,为后续优化设计提供了可靠的仿真平台和理论分析基础。

3 优化方案及实验结果

3.1 优化方案

鉴于上述沟道倾斜角导致的系统性电学偏差,本文提出了一种导通电压主动调控方案,旨在补偿由沟道倾斜角引起的阈值电压分布不均匀问题。与传统编程方案中对所有未选中字线施加固定的导通电压不同,本文方案基于目标单元(WL_n)在存储串中的垂直位置,动态调整其相邻字线(WL_{n+1}和WL_{n-1})的导通电压值,实现对阈值电压的精确补偿。

优化方案的核心机理在于充分利用了未选中字线的导通电压干扰(Pass Disturb)^[27]。具体而言,在传统编程方案中,所有未选中字线均施加相同大小的导通电压,导致不同空间位置的存储单元受到几乎一致的导通电压干扰。然而,由于沟道倾斜角的存在,存储单元的固有阈值电压沿存储串从顶部到底部逐渐增大。基于这一物理特性,本文方案创新性地提出:在对目标单元进行编程时,根据其所在垂直位置,动态调节其相邻字线的导通电压值。随着目标单元位置从顶部移向底部,逐步减小未选中单元的导通电压,每个目标字线的相邻字线都会施加一个对应的优化后导通电压(V_{OP})。通过这种渐减式的电压调控,可相应调整导通电压干扰的强度,从而有效补偿因沟道几何形变引起的阈值电压分布梯度,使整个存储串的阈值电压分布更加均匀。

该方案的具体实施架构如图6所示。由于相邻单元对目标编程单元的导通电压干扰最显著,因此本方案只对相邻单元进行导通电压调整:在编程操作过程中,仅对目标单元相邻的上下两层字线施加经过优化的导通电压,其余非相邻字线仍保持标准导通电压 V_{pass} 。

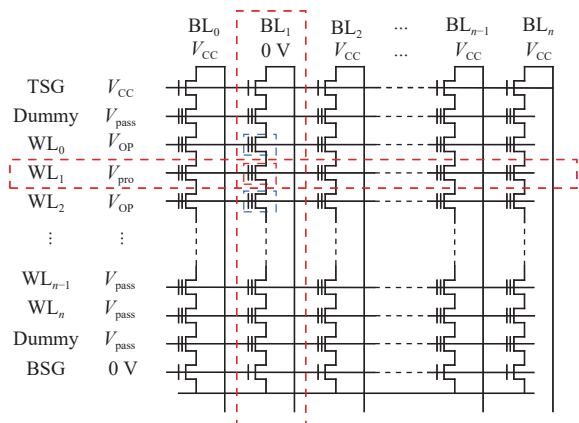


图6 闪存阵列各端口电压(改进编程方案)

Fig.6 Terminal voltages in a flash memory array (optimized programming scheme)

与传统方案相比,本文方案的创新性突出体现在3个方面:(1)针对沟道倾斜角导致的系统性电学偏差,建立了相邻单元导通电压值与目标单元空间位置的数学关系。(2)仅对目标编程单元直接相邻的两个字线进行电压调整,而非全局调整所有未选中字线,既提升了电压调控的效率,又大幅降低了算法的复杂度和硬件实现开销。(3)该方案通过纯电气方式实现阈值电压补偿,无需改变现有器件工艺结构,具有良好的向后兼容性和工程可行性。

3.2 导通电压调整方式

基于上述位置相关的导通电压调节方案,本文提出了两种具体的实现方式,分别采用线性和非线性的调整策略。这两种方法在保持核心调节机制一致的前提下,针对不同应用场景提供了灵活性选择。

线性调整方案采用固定步长的导通电压调节方式。该方案设定统一的导通电压调整步长 ΔV_{pass} ,使得相邻字线的 V_{OP} 随目标单元位置变化而线性变化,其数学表达式为

$$V_{OP}^n = V_{pass} - n \times \Delta V_{pass}$$

式中, n 为目标单元的字线层数索引(从顶部到底部递增), V_{OP}^n 为当目标单元位于WL_n时,对其相邻字线施加的 V_{OP} , V_{pass} 为基准导通电压值(固定8 V)。这种线性变化模式实现简单,调节规律明确,特别适合对算法复杂度要求较高的应用场景。

非线性调整方案则采用可变步长的调节方式。该方案根据存储串阈值电压偏移的分布特性,在不同位置采用差异化的调整步长。本文中,在阈值电压偏移较小的上半层单元(WL₀~WL₄)中,采用较小的调整步长;而在阈值电压偏移较大的下半层单元(WL₅~WL₉)中,则相应增大调整步长,强化补偿效果。这种非线性调节能够更精准地匹配锥形沟道引起的非线性阈值电压梯度,在关键区域提供更强的补偿能力,从而实现更优的整体均匀性。

值得指出的是,本文方案具有良好的可扩展性。当应用于200层以上的高层数3D NAND闪存时,只需将调整步长 ΔV_{pass} 按比例缩小至适当范围,同时保持基本调节原则不变,即可实现对更高层数器件阈值电压分布的有效补偿。这种可扩展特性使本文方案能够适应未来3D NAND技术向更高堆叠层数发展的需求。

3.3 参数确定与批次泛化

本文提出的动态导通电压调节方案旨在通过电

气方式补偿系统性工艺偏差引起的阈值电压梯度。为确保该方案在大规模生产中的可行性与经济性,基于已有的研究结果确定参数:同一制造批次内闪存芯片的阈值电压分布具有普遍的统计相关性^[28]。这一规律表明,批次内的电学变化以方向一致的系统性偏移为主,而非完全随机波动。

基于此,本文构建了一套系统性的参数确定与泛化流程:(1)对少量代表性样片进行阈值电压分布测试,建立其与导通电压的映射模型。(2)将所得模型参数应用于同批次所有芯片,实现无需逐颗校准的统一补偿。该流程将系统性的工艺偏差转化为可预测、可补偿的固定模式,有效降低了测试与部署成本。

针对高层数3D NAND器件,该方案展现出良好的可扩展性。随着堆叠层数增加,沟道刻蚀角度、氧化层厚度及沉积应力等因素导致的纵向电学梯度更为显著。本机制可依据阈值电压沿字线的空间分布规律进行灵活适配:对于连续漂移的电压梯度,可将存储串划分为若干逻辑区段并分配差异化调整步长;对于存在局部畸变的区域,则可实现逐字线级别的精细调节。将分区化补偿策略与批次统计特性相结合,能在确保芯片间性能一致性的同时,精准适配纵向结构变化。实验证明,该动态导通电压补偿机制可显著改善3D NAND的阈值电压分布均匀性。

需要指出的是,本文提出的补偿模型采用批次级参数固定化的方式。差值电压参数在出厂测试阶段通过样片拟合获得,并存储于同批芯片中,运行时保持恒定,仅依赖现有DAC设定。

现有商用DAC一般具备8 bit以上分辨率(电压步进约10~20 mV),其精度足以满足200层及以上3D NAND结构的层间电压补偿需求。由于本文方案仅在初始化阶段设定DAC输出,不涉及动态调节,因此不会增加时序开销或影响写入速度。

3.4 实验结果

3.4.1 线性调整方案

基于对存储串电气特性的系统分析,本文对线性调整方案进行了详细的仿真验证。为确保编程操作期间所有未选中单元维持导通状态,同时参考图5所示的阈值电压分布数据(其中存储串最大阈值电压为2.802 V),将最小优化后导通电压 $V_{\min}^{\text{op}}$ 设定为2.802 V。在保留足够设计裕量的前提下,选取 ΔV_{pass} 为0.4 V和0.5 V两种步长值,通过TCAD仿真平台评估了不同调整步长对阈值电压分布的优化效果。

仿真结果如图7所示,从曲线变化趋势可以清晰看出,经过优化后,阈值电压分布的整体梯度显著降低,曲线平坦度得到明显改善。原始曲线(优化前)呈

现出从 WL_0 到 WL_9 持续上升的明显梯度,调整步长 ΔV_{pass} 的取值为0.4 V时,曲线整体梯度明显降低,分布更为集中。 ΔV_{pass} 进一步取0.5 V进行优化时,曲线趋于平坦,展现出最优的均匀性改善效果。

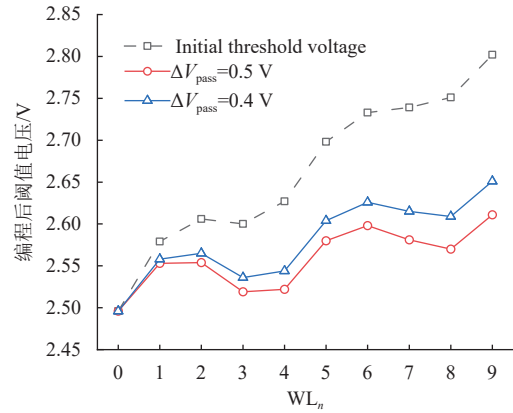


图7 线性调整方案对存储串阈值电压分布的优化效果
Fig.7 Optimization effect of linear optimization scheme on threshold voltage distribution in memory strings

具体而言,当 $\Delta V_{\text{pass}}=0.4$ V时,存储串的阈值电压最大差值 ΔV_{th} 从基准值的0.306 V显著降低至0.155 V,相对改善幅度达到49.3%。同时,阈值电压分布的标准偏差 σ_{th} 从0.092 V优化至0.045 6 V,降幅为49.8%。当 $\Delta V_{\text{pass}}=0.5$ V时,优化效果得到进一步提升: ΔV_{th} 进一步减小至0.115 V,相对于原始数据改善62.4%; σ_{th} 也降低至0.035 V,降幅达61.4%。

仿真数据充分验证了线性调整方案在补偿沟道倾斜效应方面的有效性,具体表现为:该方案能够显著降低阈值电压分布偏差,提高分布均匀性,同时确保编程操作的可靠性。特别是在 $\Delta V_{\text{pass}}=0.5$ V时,该方案在实现复杂度和优化效果之间取得了最佳平衡。此外,在整个仿真过程中,即使在最坏情况下,优化后的导通电压仍能确保高于2.802 V的安全阈值,完全满足编程操作的可靠性要求。该方案通过简洁的线性调节机制,实现了对存储串阈值电压分布的有效优化。

3.4.2 非线性调整方案

在严格保证最小导通电压安全阈值的前提下,本文对非线性调整方案进行了深入的仿真分析与性能评估,为了研究不同电压步长组合对阈值电压分布的优化效果,设置了3种分段步长配置方案,具体参数如表2所示。

仿真结果如图8所示,Case 1方案展现出最优的补偿性能,将 ΔV_{th} 降低至0.111 V,相对改善幅度达到63.7%。同时, σ_{th} 从0.092 V优化至0.032 4 V,降幅达64.3%,表现出卓越的阈值电压分布均匀性改善效果。

表 2 分段调整导通电压步长配置方案
Table 2 Optimized pass voltage tuning with step segmentation

方案	WL ₀ ~ WL ₄	WL ₅ ~ WL ₉
Case 1	0.4 V	0.6 V
Case 2	0.3 V	0.6 V
Case 3	0.3 V	0.5 V

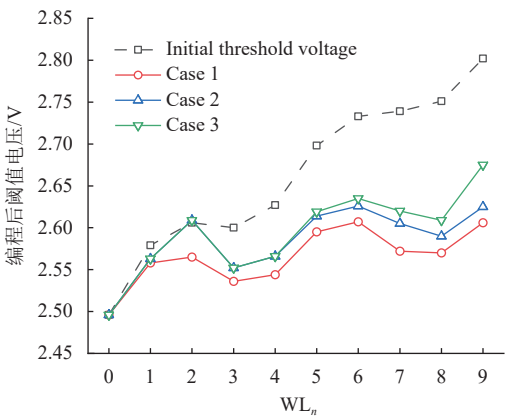


图 8 非线性调整方案对存储串阈值电压分布的优化效果
Fig.8 Optimization effect of nonlinear optimization scheme on threshold voltage distribution in memory strings

相较于线性调整方案,非线性调节方式针对阈值电压的分布特性,通过采用差异化的导通电压调整步长策略,能够更精确地匹配锥形沟道引起的非线性阈值电压梯度分布特性。本文方案中,在阈值电压偏移较大的下半层区域(WL₅ ~ WL₉)采用较大的调整步长(0.6 V),提供更强的补偿能力;而在偏移量较小的上半层区域(WL₀ ~ WL₄)则采用相对精细的调节步长(0.4 V),实现更精确的电压控制。这种分区优化的设计理念有效提升了整体补偿效果。

所有测试方案均严格满足最小导通电压的安全要求,在确保编程可靠性的同时,Case 1方案展现出的优异性能为高层数3D NAND存储器的阈值电压优化提供了技术参考。非线性调整方案通过分段式的电压分配策略,基于物理模型的有效补偿机制,实现了对沟道倾斜效应的高精度补偿。

3.4.3 相邻单元稳定性分析

在验证优化效果的同时,本文研究对动态导通电压调节方案可能产生的副作用进行了系统性评估,重点分析了相邻单元阈值电压的稳定性。仿真结果表明,采用改进编程方案后,目标单元相邻上下两层单元的阈值电压偏移量极小。如图9所示,优化前后相邻单元的阈值电压曲线几乎完全重合,呈现出高度一致的分布规律和变化趋势。具体而言,下层单元(WL₁ ~ WL₉)的阈值电压在优化后始终保持与优化前平行的变化轨迹,最大偏移量仅3 mV;而上层单

元(WL₀ ~ WL₈)的阈值电压曲线同样呈现同步变化特征,最大偏移量不超过3 mV。值得注意的是,两条曲线在整体保持高度一致性的同时,在WL₄ ~ WL₆区间显示出略微增大的波动(但仍小于3 mV),这一现象可能与该区域的电场分布特性有关。总体而言,上层和下层相邻单元的阈值电压峰值变化均被控制在3 mV以内,且整体变化趋势保持一致,这一波动完全处于器件正常工作的容差范围内。

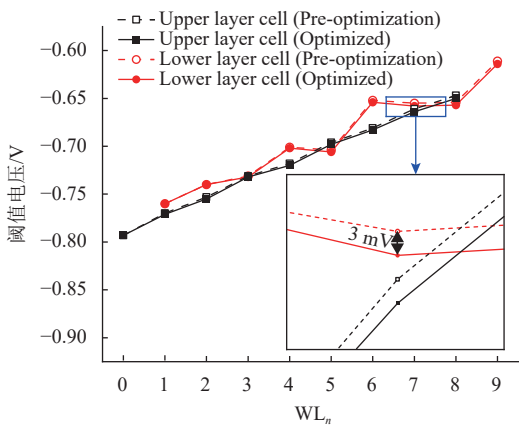


图 9 传统方案与改进方案对目标单元编程后相邻单元的阈值电压
Fig.9 Threshold voltage of adjacent cells after target cell programming with conventional and optimized schemes

表3进一步对比了目标单元的阈值电压优化幅度与相邻单元的阈值电压偏移量的关系。数据清晰显示,目标单元的优化幅度从WL₁的21 mV增加至WL₉的196 mV,而相邻单元(包括上层和下层)的偏移量始终被抑制在3 mV以下的极低水平。定量分析表明,目标单元的优化效果达到相邻单元偏移量的10~60倍,其中在WL₉处达到最大比值(196 mV/3 mV≈65倍)。这一结果充分证明,所提出的动态调节方案在显著改善目标单元阈值电压分布均匀性的同

表 3 目标单元与相邻单元阈值电压偏移量
Table 3 Threshold voltage offset in target and adjacent cells

字线层数	目标单元 偏移量/mV	上层单元 偏移量/mV	下层单元 偏移量/mV
WL ₀	0	—	0
WL ₁	21	0	0
WL ₂	41	1	1
WL ₃	64	2	1
WL ₄	83	1	2
WL ₅	103	2	2
WL ₆	126	2	3
WL ₇	167	2	3
WL ₈	181	3	3
WL ₉	196	3	—

时,对相邻单元的影响可以忽略不计。该方案通过精确定位调控的设计理念,仅对关键相邻单元进行适度电压调节,在确保优化效果的同时最大限度地降低了对非目标单元的干扰,展现了优异的可靠性和实用性。

3.4.4 写入速度影响分析

在前文验证了导通电压补偿机制对阈值电压分布及相邻单元稳定性的影响后,为进一步验证该机制对写入速度的影响,本文在 80、100、120 μs 3 种编程时间下,分别对“传统编程方案”“线性调整方案($\Delta V_{\text{pass}}=0.5\text{ V}$)”及“非线性调整方案(Case 1)”进行了阈值电压分布仿真,统计各字线的阈值电压均值与标准差,结果如表4所示。

表 4 不同编程时间下的阈值电压均值与标准差
Table 4 Mean and standard deviation of threshold voltage under different programming times

方案	编程时间/ μs	均值/V	标准差/V
传统编程方案	80	2.628	0.040
	100	2.739	0.041
	120	2.850	0.042
线性调整方案($\Delta V_{\text{pass}}=0.5\text{ V}$)	80	2.642	0.030
	100	2.753	0.031
	120	2.864	0.032
非线性调整方案(Case 1)	80	2.651	0.029
	100	2.762	0.030
	120	2.873	0.031

从表4可见,阈值电压均值随编程时间的延长而略有上升。各方案在 80~120 μs 范围内的平均阈值电压均呈相近的上升趋势,整体提升幅度约 0.22 V,三者之间的差异小于 0.02 V。这说明线性与非线性补偿的引入并未改变阈值电压随编程时间提升的速度,即编程效率保持一致,写入速度未受到显著影响。

此外,从标准差结果看,传统编程方案在 80~120 μs 之间的标准差约 0.04 V,线性与非线性方案分别为 0.03 V 左右,显示了补偿策略在保持编程速度不变的同时提升了层间一致性。整体上,线性和非线性补偿方案均未引起显著的阈值电压滞后或漂移现象。

综上分析可以得出:优化前后在不同编程时间下的阈值随时间变化速率基本一致,线性与非线性补偿机制对写入速度影响基本可以忽略,改进方案在保持编程速率稳定的同时,有效提升了阈值电压分布的均匀性。

4 结论

本文基于Sentaurus TCAD仿真平台,构建了3D

NAND存储串仿真模型,系统验证了沟道倾斜角对阈值电压分布均匀性的影响机制。通过模型证实,深孔刻蚀工艺引入的几何形变会导致存储单元沟道半径沿垂直方向呈梯度分布,进而引发阈值电压分布展宽,严重制约多级存储操作的可靠性。基于该模型验证的机理,本文创新性地提出了位置相关的动态导通电压调节方法,依据目标单元在存储串中的垂直位置,差异化调整其相邻字线的导通电压,从而有效补偿了因倾斜角引起的电学特性变异。仿真结果显示,该方案可显著改善阈值电压分布均匀性,在最优非线性调节策略下,最大阈值电压差从0.306 V降低至0.111 V,降幅达63.7%,分布标准差从0.092 V优化至0.032 4 V,降幅为64.3%。进一步的仿真结果表明,相邻单元的电压扰动始终被控制在 3 mV 以内,且在 80~120 μs 的编程时间范围内阈值电压变化相对稳定,说明该算法在不影响写入速度的前提下显著提升了层间阈值电压的均匀性。

本文所提出的方法为高层数3D NAND存储器提供了一种无需改变工艺、低成本且可扩展的电路级优化方案。同时,也证明了利用电学补偿策略来抵消3D NAND芯片制造中的结构误差是一条切实可行的技术路径。随着堆叠层数增加与单元间距缩小,电压调控策略在三维阵列中的串扰特性及其与现有耦合机制的交互影响仍需深入研究。后续工作可重点评估该策略在高层数(如200层以上)存储阵列中的适用性,系统分析其横向串扰效应及耦合作用,并为QLC/PLC等高密度存储技术开发更系统的电路级优化方案。

参考文献:

[1] GODA A. 3-D NAND technology achievements and future scaling perspectives[J]. *IEEE Transactions on Electron Devices*, 2020, 67(4): 1373-1381.

[2] SONG J, LUO X, TANG Y, *et al.* On-chip optical sense strategy illustrated in the case of flash memory system[C] // 2016 IEEE International Electron Devices Meeting (IEDM) . San Francisco: IEEE, 2016: 1-22.

[3] KANG H J, CHOI N, BAE J H, *et al.* Analysis of clockwise and counter-clockwise hysteresis characteristics in 3-D NAND flash memory cells[J]. *IEEE Electron Device Letters*, 2017, 38(7): 867-870.

[4] PARK G H, CHO W J. Reliability of modified tunneling barriers for high performance nonvolatile charge trap flash memory application[J]. *Applied Physics Letters*, 2010, 96(4): 043501.

[5] BHATT U M, MANHAS S K, KUMAR A, *et al.* Mitigating the impact of channel tapering in vertical channel 3-D

- NAND[J]. *IEEE Transactions on Electron Devices*, 2020, 67(3): 929-936.
- [6] LEE J G, JUNG W J, PARK J H, *et al.* Effect of the blocking oxide layer with asymmetric taper angles in 3-D NAND flash memories[J]. *IEEE Journal of the Electron Devices Society*, 2021, 9: 774-777.
- [7] NOWAK E. Intrinsic fluctuations in vertical NAND flash memories[C] // 2012 Symposium on VLSI Technology (VLSIT) . Honolulu: IEEE, 2012: 21-22.
- [8] KIM K T, AN S W, JUNG H S, *et al.* The effects of taper-angle on the electrical characteristics of vertical NAND flash memories[J]. *IEEE Electron Device Letters*, 2017, 38(10): 1375-1378.
- [9] KIM H, AHN S J, SHIN Y G, *et al.* Evolution of NAND flash memory: from 2D to 3D as a storage market leader[C]// IEEE International Memory Workshop (IMW) . Monterey: IEEE, 2017: 1-4.
- [10] OH Y T. Impact of etch angles on cell characteristics in 3D NAND flash memory[J]. *Microelectronics Journal*, 2018, 79: 1-6.
- [11] LU C C, CHENG C C, CHIU H P, *et al.* Analysis and realization of TLC or even QLC operation with a high performance multi-times verify scheme in 3D NAND flash memory[C]// 2018 IEEE International Electron Devices Meeting (IEDM) . San Francisco: IEEE, 2018: 1-2.
- [12] ZHAO C L, LI D. Investigation of threshold voltage distribution temperature dependence in 3D NAND flash[J]. *IEEE Electron Device Letters*, 2018, 40(2): 204-207.
- [13] LI H. Modeling of threshold voltage distribution in NAND flash memory: a Monte Carlo method[J]. *IEEE Transactions on Electron Devices*, 2016, 63(9): 3527-3532.
- [14] LEE D H, SUNG W. Decision directed estimation of threshold voltage distribution in NAND flash memory[J]. *IEEE Transactions on Signal Processing*, 2014, 62(4): 919-927.
- [15] CHUN M, LEE J, KIM M, *et al.* RiF: improving read performance of modern SSDs using an on-die early-retry engine[C]// 2024 IEEE International Symposium on High-Performance Computer Architecture (HPCA) . Edinburgh: IEEE, 2024: 643-656.
- [16] PIAO Z, WEI D, XIANG H, *et al.* LVDE: a lightweight threshold voltage distribution estimation strategy for high-performance 3-D NAND flash memory[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2025, 72(9): 4723-4736.
- [17] LIAO H Y, HSU W L, HSIEH J W, *et al.* Read retry mechanism for 3D NAND flash memory: observations, analyses, and solutions[C]// 2024 13th Non-Volatile Memory Systems and Applications Symposium (NVMSA) . Taipei: IEEE, 2024: 1-6.
- [18] 罗铮, 韩国军. 3D NAND闪存P/E次数的快速测评与寿命预测[J]. *广东工业大学学报*, 2024, 41(6): 52-59.
- LUO Z, HAN G J. Rapid measurement and lifetime prediction of 3D NAND flash P/E cycles[J]. *Journal of Guangdong University of Technology*, 2024, 41(6): 52-59.
- [19] 李应钊, 朱广平, 韩国军. 3D NAND闪存多维度噪声耦合错误检测及信道建模[J]. *广东工业大学学报*, 2025, 42(6): 1-9.
- LI Y Z, ZHU G P, HAN G J. Multi-dimensional noise coupling error detection and channel modeling for 3D NAND flash memory[J]. *Journal of Guangdong University of Technology*, 2025, 42(6): 1-9.
- [20] MACONI A, ARREGHINI A, COMPAGNONI C M, *et al.* Comprehensive investigation of the impact of lateral charge migration on retention performance of planar and 3D SONOS devices[J]. *Solid-State Electronics*, 2012, 74: 64-70.
- [21] TANAKA H. Bit cost scalable technology with punch and plug process for ultra high density flash memory[C]// 2007 Symposium on VLSI Technology (VLSIT) . Kyoto: IEEE, 2007: 14-15.
- [22] JO H, AHN S, SHIN H. Investigation and modeling of Z-interference in poly-Si channel-based 3-D NAND flash memories[J]. *IEEE Transactions on Electron Devices*, 2022, 69(2): 543-548.
- [23] AHN S. An innovative program scheme for reducing Z-interference in charge-trap-based 3-D NAND flash memory[J]. *IEEE Transactions on Electron Devices*, 2023, 70(12): 6695-6698.
- [24] MITROFANOV O, MANFRA M. Poole-Frenkel electron emission from the traps in AlGaIn/GaN transistors[J]. *Journal of Applied Physics*, 2004, 95(11): 6414-6419.
- [25] LUI O K B, MIGLIORATO P. A new generation-recombination model for device simulation including the Poole-Frenkel effect and phonon-assisted tunnelling[J]. *Solid-State Electronics*, 1997, 41(4): 575-583.
- [26] LEE J D, LEE C K, LEE M W, *et al.* A new programming disturbance phenomenon in NAND flash memory by source/drain hot-electrons generated by GIDL current[C]// 2006 21st IEEE Non-Volatile Semiconductor Memory Workshop. Monterey: IEEE, 2006: 31-33.
- [27] LEE J D, HUR S H, CHOI J D. Effects of floating-gate interference on NAND flash memory cell operation[J]. *IEEE Electron Device Letters*, 2002, 23(5): 264-266.
- [28] LI Q, ZHOU Y, CHEN S, *et al.* Shaving retries with sentinels for fast read over high-density 3D flash[C]// 2020 53rd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO) . Athens: ACM, 2020: 483-495.

(责任编辑: 杨耀辉 英文审核: 熊荣斌)