

一种基于位线电荷循环的低功耗SRAM阵列设计

张瀚尊¹ 贾嵩^{2,†} 杨建成¹ 王源²

1. 北京大学微电子学研究所, 北京 100871; 2. 北京大学微电子器件与电路重点实验室, 北京 100871;

† 通信作者, E-mail: jias@pku.edu.cn

摘要 为了降低静态随机存储器(SRAM)的动态功耗, 提出一种基于位线电荷循环的读写辅助电路的 SRAM 阵列。与传统设计相比, 辅助电路中转和保存了在读写操作中本该被直接泄放掉的位线电荷, 并重新用于下一个周期的位线充电。提出的 SRAM 存储器采用标准 14 nm FinFET spice 模型搭建, 电源供电电压为 0.8 V。仿真结果表明, 与传统设计相比, 提出的存储阵列的功耗可以降低 23%~43%, 并将 SNM 和 WNM 至少提高 25% 和 647.9%。

关键词 SRAM; 位线电荷循环; 读写辅助

A Charge Recycling Scheme with Read and Write Assist for Low Power SRAM Design

ZHANG Hanzun¹, JIA Song^{2,†}, YANG Jiancheng¹, WANG Yuan²

1. Institute of Microelectronics, Peking University, Beijing 100871; 2. Key Laboratory of Microelectronic Devices and Circuits, Peking University, Beijing 100871; † Corresponding author, E-mail: jias@pku.edu.cn

Abstract In order to cut down the dynamic power of static random access memory (SRAM), a bitline charge cycling based read and write assist circuit for SRAM is presented. Compared with the traditional design, the assist circuit saves and reuses the bitline charge which should be directly discharged during read and write operation to reduce bitlines charging power consumption in the next cycle. The SRAM memory is built by the SMIC 14 nm FinFET spice model, and the power supply voltage is 0.8 V. The simulation results show that the power consumption of the proposed SRAM array is reduced by 23%–43% compared with the traditional design, and the SNM and WNM has increased by at least 25% and 647.9% respectively.

Key words SRAM; bitline charge cycling; read and write assist

目前, SOC 芯片离不开基于 SRAM 的高速缓存, 而 SRAM 往往占据 SOC 的绝大部分面积, 因此低功耗 SRAM 对移动设备 SOC^[1]非常重要。近年来, 尽管 FinFET 工艺以更小的晶体管尺寸和更低的功耗广泛应用于 SRAM 的制作, 但更低的工作电压和相对剧烈的工艺波动使其需要搭配使用读或写辅助电路才能实现预期的读写指标^[2]。辅助电路会带来额外的能量损失, 导致 SRAM 阵列的功耗不能随着电源电压的下降而显著降低。

在 SRAM 阵列中, 动态功耗主要来自对大扇出位线进行预充电或电压状态切换^[3]。例如, 在采用

标准 14 nm 工艺 spice 模型搭建的位线全摆幅阵列仿真中, 位线预充电消耗的能量占总动态功耗的 90% 以上。在常规的 SRAM 阵列的读操作结束后, 这部分位线预充电荷会被单元的存“0”节点直接泄放到 GND, 导致潜在的电能浪费。如果这部分电荷可以伴随读写操作被回收或再利用^[4-7], 则可以大幅度降低总体的动态能耗。Kim 等^[4]首次提出一种使用 0.13 μm 工艺, 电源电压为 1.5 V 的基于位线电荷循环的 SRAM 阵列, 电荷随着连续的写操作, 从第一列的位线向后面的位线传播, 实现电荷共享。这种循环方式以降低写操作的位线摆幅为代价, 与没

选择晶体管的尺寸,以保证 6T 单元功能。由于 FinFET 工艺的 Fin 只能是离散的值,为了实现高密度,FinFET 单元的尺寸没有太多的选择^[8]。由于 SRAM 单元存在读干扰和半选择问题,因此在设计尺寸时增大 SNM,并且较大的 SNM 意味着 FSBC 和 HSBC 具有较强的数据稳定性。如图 1(b)所示, β (PD 的 Fin 数:PG 的 Fin 数)越大,单元的 SNM 越大^[9]。尽管具有相同的 β 数值,但是 PU:PG:PD=1:1:1 单元的 SNM 大于 1:2:2 的单元,原因是 1:2:2 的单元具有较大的读电流和读出速度,大电流导致 PD 分压变大,从而存储稳定性较差^[10]。

在 PU:PG:PD=1:1:1, 1:2:1 和 2:1:1 条件下,测得单元写噪声容限(WNM)的变化情况,如图 1(c)所示。通常具有较大 γ (PG:PU)的单元,在位线的作用下容易翻转。从图中可以看出,尺寸比例为 2:1:1 的单元写操作失败,这是由于 PU 的驱动能力远超 PG,导致单元存储接点的状态“1”不容易被下拉成“0”,所以该比例在设计中不可取。1:2:1 的单元 WNM 最大,但是相应的 SNM 最小。

考虑到 HSBC 的存储稳定性和电路设计的复杂性,通常在 SRAM 单元尺寸选择中尽量提高单元的 SNM,对 WNM 使用写辅助电路来弥补,使设计能够满足既定的性能要求。本文仅对比例为 1:1:2 的 SRAM 阵列进行研究和优化,但提出的读写辅助方法可以应用于其他尺寸设计(如 PU:PG:PD=1:1:1)中。

1.2 电路设计的整体布置

本设计中的 256×64 SRAM 阵列基于标准 14 nm spice 模型搭建,正常工作频率为 1 GHz,具有读写辅助电路,字长为 64 bit。图 2 展示整个阵列的结构。基于位线电荷循环的读写辅助电路(charge sharing circuit, CRC)主要由两部分组成:1) 由 NVSS_en 控制的 VSS_switch,作用是确定 CVSS 是连接到 GND 还是 CS_point; 2) charge_share_part (CSP),一个负电压发生控制器。在读写操作期间,位线上的电荷通过 FSBC (本文中阵列在读写操作中没有半选单元)的 PD 和 VSS_switch 被 CSP 收集,这些电荷将在下一个读写周期到来时,重新用于位线的预充电,而不是被释放掉后再使用电源对位线进行预充电。CSP 将读辅助与写辅助结合在一起,单元负接地电压(NVSS)用于加速读操作,并提高读稳定性,而 NBL 用于写辅助,受写使能信号控制。

由于阵列中行数很多,并且会有大量漏电涌向

负电压节点,导致生成的 CS_point 的负电压减小,因此,每行 VSS_switch 中 MV 管的尺寸需要精细选择。特别地,当电路中存在负电压时,相关的晶体管不会完全关闭,这一问题可以通过调整时序来解决。

1.3 读操作和位线充电循环工作流程

图 3 为读操作的波形。首先,将 FSBC 的两根位线预充电至高电平,CSP 的初始化开始:Ini_nc 和 iso_nc 各自打开 MC 和 MI,结果使 NV_point 和 CS_point 的电压下降到 0。然后,WL 和 NVSS_en 导通,从而通过 PD 和 VSS_switch,在 CVSS 和 CS_point 之间建立连接。同时,ini_set 的信号从低电平变为高电平,由于 NV_cap 和 CS_cap 的两个极板之间的电压差不能立即改变,因此 CS_point 和 NV_point 的电压都下降到负值。如图 2(a)所示,来自存储“0”的 Q 节点相应位线的放电电荷被 CS_cap 接收,致使 CS_point 的电压从负电平上升。当 WL 关闭时,CS_point 电平接近 0。

使用 NVSS 作为读辅助有两个明显的优点:一是 CS_cap 的负电压使得 FSBC 的 CVSS 电压变为负值,增加了对应位线的放电速率,加快读出速度;二是因为 PD 一直工作在线性区,存“0”节点的电压也被拉至负电平,由此存“0”节点的稳定性得到增强。另外,NVSS 还允许位线的电平快速降低至 0,进一步减轻位线的正电平对存“0”节点的影响。存“0”节点的负电压使存“1”节点的 NMOS 被完全关断,对应的 PMOS 被加强,使得存“1”节点的稳定性更强。因此,使用 NVSS 的 FSBC 的读稳定性得到增强。

灵敏放大器(sensitive amplifier, SA)从 FSBC 读出数据后,WL 和 NVSS_en 被关断,存“0”节点的电压恢复到 0。同时,信号 ini_set 从 0 跳变为高电平,使 CS_point 的电压从低电平升至 0.7~0.8 V。当 NV_cap 和 CS_cap 的极板电压都在 ini_set 的作用下完全反转后,控制信号会控制 CS_cap 与已经被存“0”节点放电的位线电荷共享,使位线电平被从 0 抬高,于是位线在下一个读周期中的预充电能量损耗减少,最大幅度可减少 50% 的位线预充功耗。在将 NV_cap 和 CS_cap 的电平从低电压拉到高电压时,也存在翻转功耗,可将电容状态翻转的功率损耗视为电容器充电能量和电容器势能变化之和。与减少的位线预充电能量相比,这种程序的损耗是可以接受的。在传统阵列中,位线上的电荷被无意义地

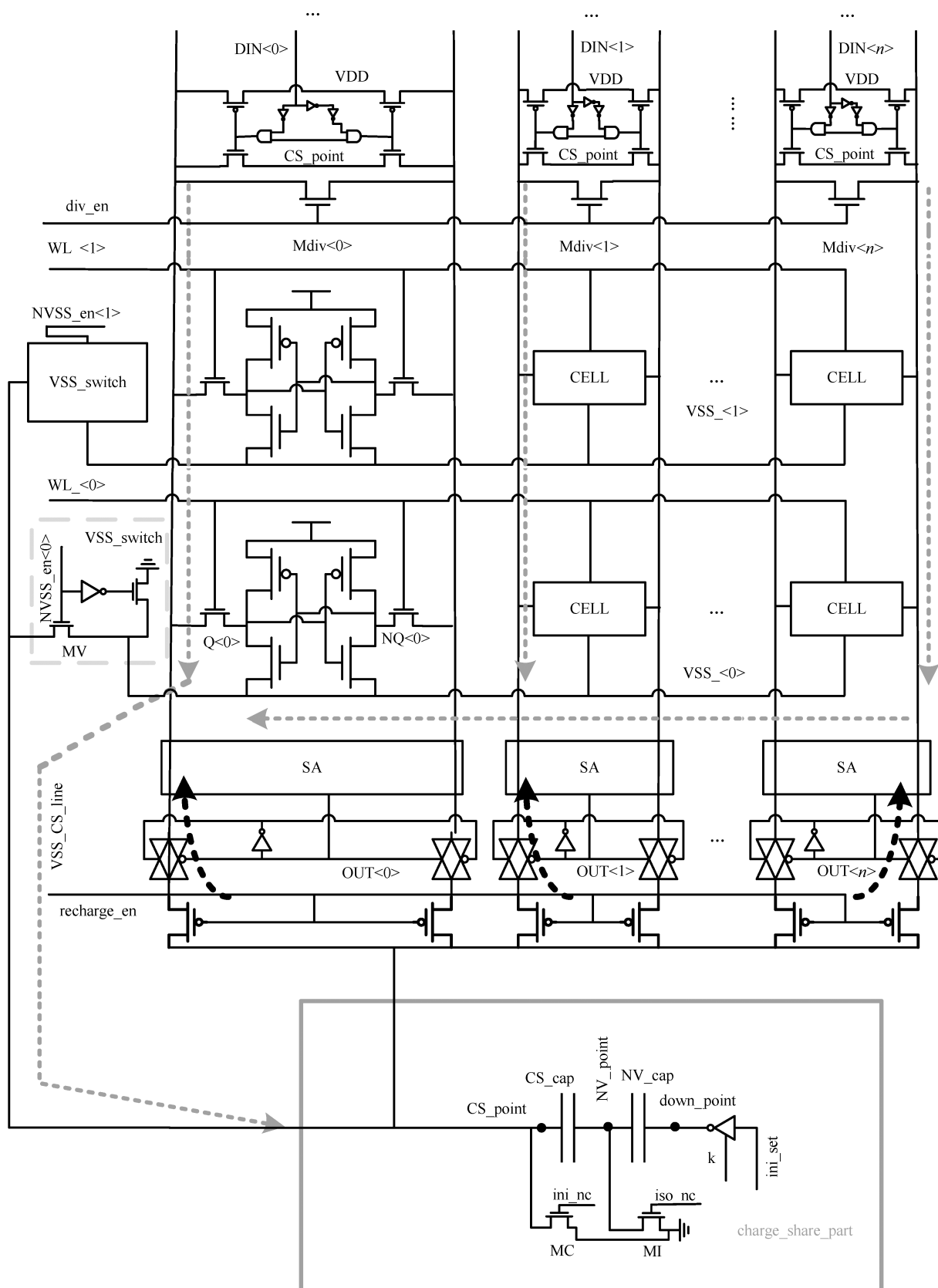


图 2 实际电路图
Fig. 2 Actual circuit diagram

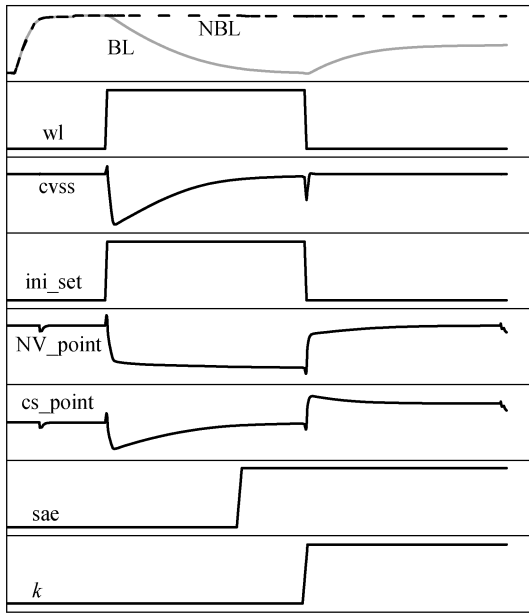


图 3 读操作的波形

Fig. 3 Waveform of read operation

放电到 GND, 但在本设计中, CSP 循环利用位线电荷有效地节省了预充电的电能耗耗。

NV_cap 和 CS_cap 的电容可通过以下公式并通过仿真来选择, 以便限制实际电路面积。

$$\begin{aligned} C_{BL} \times V_{BL} + C_{CS_point} \times V_{NV} \\ = (C_{BL} + C_{CS_point}) \times V_{itrl}, \end{aligned} \quad (1)$$

$$C_{CS_point} \times V_{recharge} = (C_{BL} + C_{CS_point}) \times V_{div}, \quad (2)$$

其中, V_{NV} 是 CSP 生成的负电压值, 几乎等于 $-0.8V$; V_{itrl} 指位线与 CS_point 做电荷共享后的平均电压; $V_{recharge}$ 是 ini_set 从高变低时 CS-point 的电压; V_{div} 表示电荷循环后的位线电压, 在 $VDD/2$ 左右。设计中, 电荷共享电容 CS_cap 略大于位线电容, 以便在

WL 脉冲结束时数据存“0”节点的位线电压可以快速降至 0 ($V_{itrl}=0$) 以及中和其余单元向负压节点的漏电。负压电容 NV_cap 至少是 CS_cap 的 4 倍, 因为 NV_cap 的意义是为 CS_cap 提供尽可能稳定的负电压, 以便 CS_cap 准备好从位线接收电荷。

1.4 写辅助电路工作流程

在写操作期间, 写辅助 NBL 由 CSP 产生的负电压生成。当写使能信号到达时, 写驱动器会控制相应的位线充电至 VDD 或与 CS_cap 共享电荷。写操作期间, CSP 无论在结构上还是在时序上都与读操作没有任何区别。WL 关闭后, 由 div_en 控制的 NMOS 对两个位线的电压进行平均。要实现的效果是, 在一个写周期之后, 在 CSP 与位线之间的电荷共享的帮助下, 两个位线的平均电压相等且不低于 $VDD/2$, 这样可以使下一个写周期开始的位线充电能耗尽量降低。

图 4 为本设计中的写驱动电路。由写输入 DIN0、写使能 W_EN 和初始化 ini_nc 这 3 个信号共同决定对相应位线进行充电还是放电到 0, 再加负载电平。

2 结果分析

本文中, SNM 和 WNM 是量化 SRAM 读写稳定度的主要指标。如图 5 所示, 使用辅助电路的 SNM 和 WNM 比不使用辅助设计的分别提高 32.6% 和 647.9%。SNM 改善的原因是 NVSS 读辅助产生带负电平的存“0”节点, 这样的“0”节点数据不容易被较高的位线或其他外部噪声电压干扰或改变, 同时存“1”节点的稳定性也得到增强。

WNM 提高的原因是使用了 NBL 技术。位线上的负电压可以使最初存储“1”的节点状态迅速翻转为“0”。可以将 WNM 近似地视为单元的翻转电压

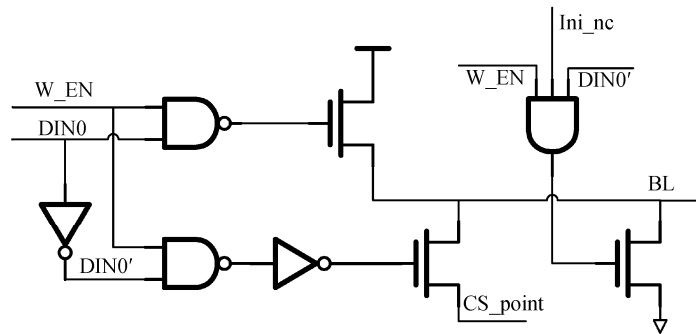


图 4 设计中采用的写驱动电路

Fig. 4 Write drive circuit used in the design

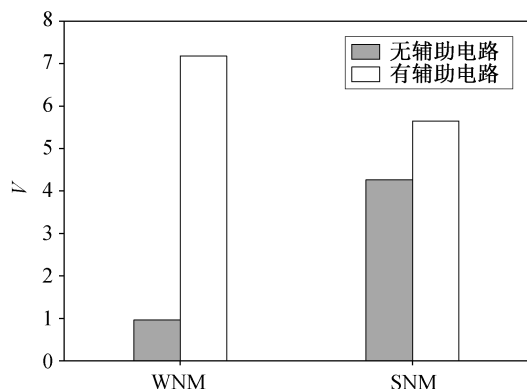


图 5 有辅助电路和无辅助电路的阵列 SNM 和 WNM 比较
Fig. 5 Comparison of array SNM and WNM with and without asstt circuit

与 VDD 之间的电压差, 而负电压可以有效地降低单元的翻转电压, 即 WNM 得到显著改善。

图 6 为有辅助与无辅助电路各部分的功耗对比情况。 $R_{w/o}$ 表示有辅助与无辅助的读操作, $W_{w/o}$ 表示有辅助与无辅助的写操作。实际上, 读操作仅使预充电的位线放电, 因此 R_w 和 R_o 之间的读操作单元的功耗没有显著的差异。在使用 NBL 辅助的情况下, 该单元的写功耗比未使用 NBL 降低约 43%。这是因为 NBL 增加了写速度, 使得 FSBC 在两个稳定状态之间切换的时间大大减少, 即减少了 VDD 与 GND 之间切换电流的导通时间, 从而显著地降低用于位单元写操作的写功耗。

WL 关闭之后, 通过 CS_{cap} 与位线之间的的电荷共享, 使位线的电压被上拉, 从而在下一个周期开始时降低预充电功耗。如图 5 所示, 读写周期中的预充电或电压转换功耗分别降低 50% 和 68.3%, 与预充电功耗的降低相比, 辅助电路的功耗 CRC

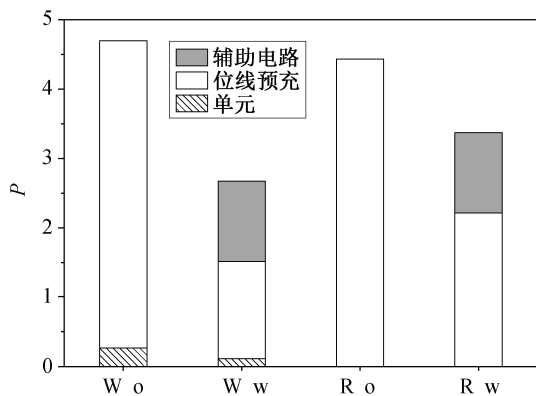


图 6 阵列不同部分的能耗对比

Fig. 6 Energy consumption comparison of different parts

可以忽略。CRC 电路的功耗主要包括两个电容电压切换的损耗以及电容器电势能的变化(NV_{cap} 电容器的电荷在第一次开机时要消耗大量电能来完成充电, 以便起到提供稳定负压的作用)。

在常规阵列的读操作中, 从 VDD 获得的功率是用来将位线预充电至 VDD, 单元本身不会消耗太多的能量。并且, 只要 SA 有较好的灵敏度和 PVT 包容度, 没有读辅助功能的 SRAM 设计也不需要全幅位线。在常规阵列的写操作中, 根据输入对大扇出位线电压进行切换是写操作的主要动态功耗来源。在本文的阵列中, CRC 用于实现位线从 VDD 到 GND 的完全放电, 并通过充电循环, 将其重新充电回 VDD/2。因此, 在这种设计中, 位线与 CRC 之间有相对大量的电荷流动, 与没有辅助的设计相比, 本文的设计更容易加热。如图 7 所示, 热功率 heat_{pwr} 明显大于单纯来自外部电源 vdd_{pwr} 的功率。

图 8 为具有辅助电路和不具有读操作的位线电

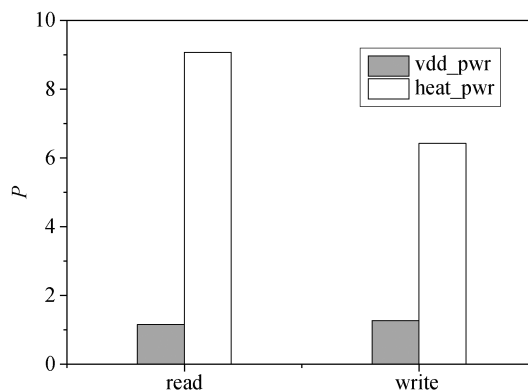


图 7 电源的能耗和热功耗的对比

Fig. 7 Comparison of power consumption and thermal power consumption

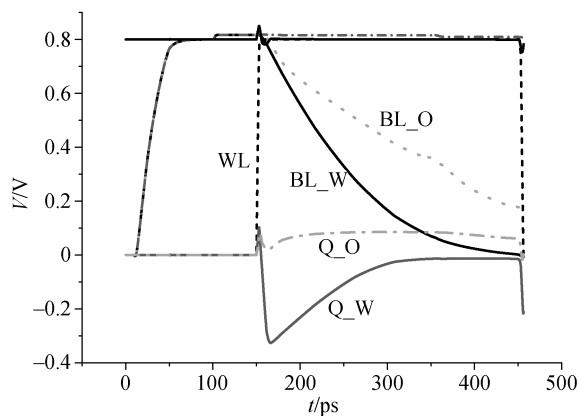


图 8 读操作位线在有辅助电路和无辅助电路的波形对比

Fig. 8 Voltage waveform comparison of the bit line with and without assist circuit

压变化的比较。没有辅助电路的位单元的位线电压不能在 300 ps 内降至 0, 而 NVSS 读辅助可改善这一情况。在本设计中, 数据“0”节点的位线可以在读过程中借助 CRC 完全放电, 因此 SA 对 FinFET 的 PVT 波动不大敏感。

在每个读写周期结束时, CS_cap 和位线进行由 CRC 控制的充电循环, 减少下一个周期中位线的预充电能量。但是, 此过程将需要一定的时间才能完成, 因此结构将占用一个周期中更多的时间, 降低了工作频率。

图 9 为基于不同读写辅助电路的 SRAM 读写功耗对比情况, 各个电路在 1 GHz 工作频率下的最小工作电压都可以达到 0.6 V。通常写辅助电路会增加 SRAM 的写入功耗, 如文献[5-6,8]中的写入功耗分别为不使用辅助电路的 1.59, 1.28 和 1.46 倍, 这

是由于文献[5-6]中用于电荷共享的电荷被直接放掉了, 文献[8]中的电荷循环利用率并不高, 并且辅助效果也不理想。本文提出的辅助电路使用额外的电能来改变共享电容的电压状态, 维护电荷循环的稳定性, 并且可以显著地减少位线的状态变化功耗, 所以总体功耗比不使用辅助电路的 SRAM 阵列的写功耗更小, 仅为 58%。

在一个读周期内, 放电的位线会在周期末充电至高电平, 所以读周期的能耗主要是位线充电的功耗, 与不用辅助电路的 SRAM 相比, 读功耗下降 23%。

对 SRAM 来说, 还需要考虑电路的工艺、电压和温度等 PVT 因素对其读写特性的影响, 图 10 为本设计读操作的蒙特卡洛仿真结果, 可以看出, 在不同工艺角和温度下, 在位线电压在辅助电路的帮

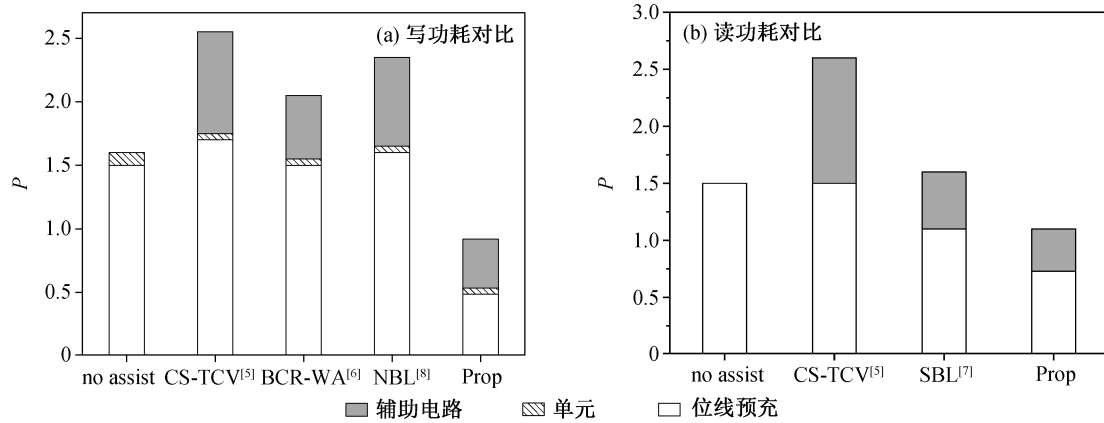
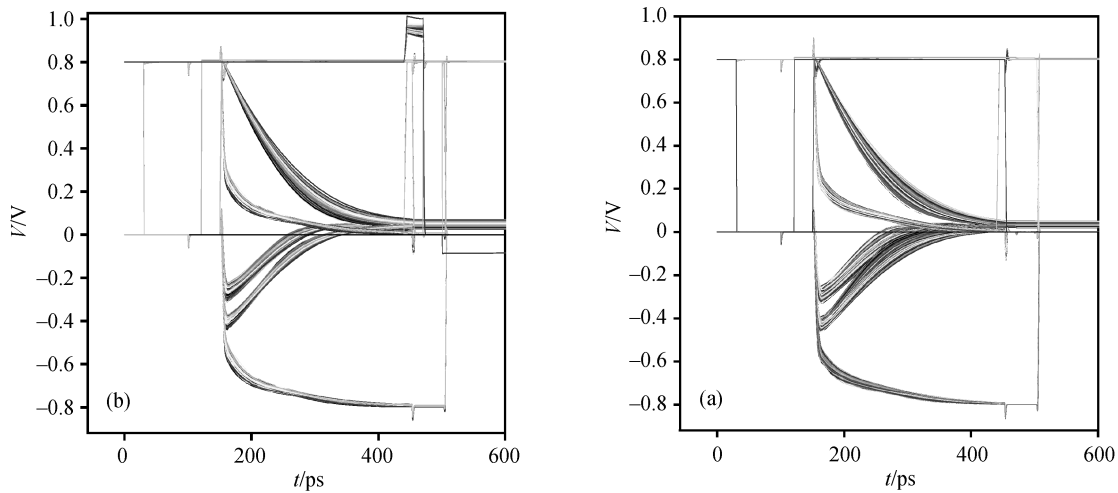


图 9 不同辅助电路的能耗对比(工作电压为 0.8 V)

Fig. 9 Power consumption comparison of different assist circuits (VDD is 0.8 V)



(a) 不同工艺角下的波形; (b) 25°C 和 125°C 下的波形

图 10 读操作的 2000 点 Monte Carlo 仿真结果

Fig. 10 2000-point Monte Carlo simulation result of read operation

助下, 300 ps 内可以降至 200 mV 以下, 从而可以使 SA 稳定地工作。在 25°C 和 125°C 条件下, 位线电压的区别不大, 都可以在 300 ps 内降至 0。因此, 本设计表现出较好的抗 PVT 特性。

3 结论

本文提出一种具有基于位线电荷循环思想的读写辅助电路的 SRAM 阵列, 旨在获得更低的功耗、更快的读出速度以及更好的 SNM/WNM。基于标准 14 nm spice 模型的 0.8V SRAM 阵列仿真结果显示, 与传统的 SRAM 阵列相比, 辅助电路可以将总功耗降低 23%~43%, 并有效地将 SNM 和 WNM 分别提高 25% 和 647.9%。

参考文献

- [1] Yang B. A low-power SRAM using bit-line charge-recycling for read and write operations. *IEEE Journal of Solid-State Circuits*, 2010, 45(10): 2173–2183
- [2] Jeong H, Kim T, Yang Y, et al. Offset-compensated cross-coupled PFET bit-line conditioning and selective negative bit-line write assist for high-density low-power SRAM. *IEEE Transactions on Circuits & Systems I: Regular Papers*, 2015, 62(4): 1062–1070
- [3] Chen Y H, Chan W M, Wu W C, et al. A 16 nm 128 Mb SRAM in high- κ metal-gate FinFET technology with write-assist circuitry for low-VMIN applications // 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers. San Francisco, 2014: 238–239
- [4] Kim K, Mahmoodi H, Roy K. A low-power SRAM using bit-line charge-recycling. *IEEE Journal of Solid-State Circuits*, 2008, 43(2): 446–459
- [5] Karl E, Guo Z, Conary J W, et al. A 0.6 V 1.5 GHz 84 Mb SRAM design in 14 nm FinFET CMOS technology // 2015 IEEE International Solid-State Circuits Conference, 2015, 51(1): 222–229
- [6] Jeong H, Oh S H, Oh T W, et al. Bitline charge-recycling SRAM write assist circuitry for V_{\min} improvement and energy saving. *IEEE Journal of Solid-State Circuits*, 2019, 54(3): 896–906
- [7] Choi W, Park J. A charge-recycling assist technique for reliable and low power SRAM design. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2016, 63(8): 1164–1175
- [8] Song T, Rim W, Park S, et al. A 10 nm FinFET 128 Mb SRAM with assist adjustment system for power, performance, and area optimization. *IEEE Journal of Solid-State Circuits*, 2017, 52(1): 240–249
- [9] Aburahma M H, Anis M. Nanometer variation-tolerant SRAM. New York: Springer, 2013
- [10] Wang A, Calhoun B H, Chandrakasan A P. Sub-threshold design for ultra low-power systems. New York: Springer, 2006