题记：自己曾一度被此问题所困扰，去论坛搜索网友讨论及向老师同学请教后，虽仍不能完全了然于心，但终有所悟，仓促间整理出这篇文字以作记录。因水平所限，文中定有不肯之处，错误所在还请前辈多多指教。或许这个问题对高手来说不屑一提，但还是谨以此文献给像我一样常常困惑于模拟集成电路设计诸多未知及疑问的新手和同学们。

Sumig 于 2008 年 10 月末

在 Spectre 仿真中，对于 Vdsat 和 Vov 不相等的初次探究

在进行仿真时，常常会发现这样一个问题，那就是$V_{ov} = V_{gs} - V_{th} > V_{dsat}$

vds	344.043m
vdsat	244.943m
vgb	1.10365
vgd	759.611m
vgs	1.10365
vth	790.913m

正如左图所示 NMOS 的工作状态，可以看出：

$$V_{ov} = V_{gs} - V_{th} = 1.10365 - 0.190913$$
$$= 0.3045587V > V_{dsat} = 0.244943V$$

那么是什么造成了这种现象，我们如何从中判断 MOS 管是否饱和呢 ?

下面本文将对这个问题进行初步试探性的分析，简便起见仅以 NMOS 为例。

对于一个普通的 NMOS 器件，在只考虑沟道长度调制效应时，其饱和条件我们知道是 $V_{ov} = V_{gs} - V_{th} \le V_{ds}$，此时导电沟道被夹断，我们得到一个更加精确的 I-V 公式：

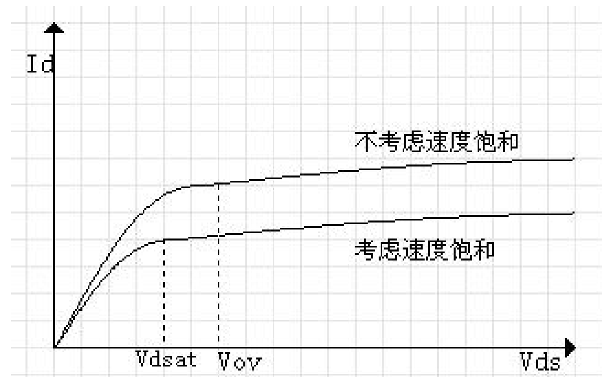
$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{gs} - V_{th})^2 \cdot (1 + \lambda V_{ds})$$

此条件下，$V_{ov} = V_{gs} - V_{th} = V_{dssat}$ 是 NMOS 的饱和临界点。

然而在实际情况中，短沟道器件存在着速度饱和效应，当 MOS 沟道很短时，沟道内的电场强度会变得很强，场强达到一定程度时，载流子速度达到最大值，从而使得沟道电流也达到饱和。

当我们考虑短沟道的这一效应时，在给定已知的 Vov 下，可能会发生这种情况：在 Vds 还没有上升到 Vov 时（沟道尚未夹断），NMOS 就因为速度饱和效应而提前进入饱和态。这种提前饱和的现象导致漏电流要比不发生速度饱和时的电流值小。

这里把 MOS 管恰好进入饱和态时 Vds 定义为饱和临界点 Vdsat，这个值是小于 Vov 的。Vdsat 和 Vov 如右图所示。



在 Cadence Model Manual BSIM3v3

有 $V_{dsat} = \dfrac{V_{gs} - V_{th}}{A_{bulk}}$

或者$V_{dsat} = \dfrac{V_{gs} - V_{th}}{L \cdot E_{sat}}$

Razavi 教材中有 MOS 速度饱和时的 I_D 详细解析式，具体内容参考 Razavi P480。

或许我们可以这样认为，在 Spectre 仿真后，如果发现 Vdsat 比 Vov 小很多，那么就可以断定 MOS 管发生了速度饱和，此时决定饱和态的临界点不再是 Vov，而是 Vdsat。只要 $V_{ds} \geq V_{dsat}$，那么便可以认为 MOS 进入了饱和态。

然而在我们是不期望发生速度饱和效应的，因为它减小了漏电流的值，使 Vov 的驱动效率降低。所以在设计参数时，我们需要注意取合理的 Vds，不致沟道场强过强。

Razavi 指出，当场强达到 $1V/\mu m$，载流子的迁移率开始下降，场强进一步曾前的话，载流子可能会达到速度饱和。根据这一结论，我们或许会找到一个初步的不甚精确的 Vds 临界值。下面以 SMIC 0.18um 工艺为例，计算在沟道长度取最小尺寸时，NMOS 发生速度饱和效应的 Vds 临界值。

n18：$L = 180nm \quad E = 1V/\mu m$ 则 $V_{ds,\max} = E \cdot L = 0.18V$

n33：$L = 350nm \quad E = 1V/\mu m$ 则 $V_{ds,\max} = E \cdot L = 0.35V$

从上面分析可以看出，用增大沟道长度的方式能够延缓速度饱和效应的发生，但是着意味着更大的寄生效应以及阈值电压 Vth 随之变大。

关于沟道长度调制效应、Vth 随 L 变化以及速度饱和效应，level2 到 level49 的模型都给予了相应的合理考虑，对此本文不再一一赘述，如有兴趣，可以参阅相关文档以加深理解。

参考文献：

1. Razavi 第 2 章、第 16 章
2. www.edaboard.com　Forum
3. 最后附上网友的相关讨论，以供大家参考，同时也欢迎大家进一步讨论

==

Question:

Why vdsat is not equal to vgs-vth?

I have simulated some circuits using Hspice.

But seeing the list file, vdsat is not equal to vgs-vth like below.

vgs 957.6501m 876.6040m

vth 726.1315m 635.6738m

vds 2.5404 342.4499m

vdsat 191.1505m 195.5912m

Answer:

- Because vdsat is not equal to vgs-vth especially for short-channel devices.

I guess generally, they are v.close....

The problem is you need to get a point where your switching of liner to sat. It is happening and accurate W/L for that may not be applicable in spice/layout.

Also vdsat doesn't always remain vgs-vth due to lamda effect..., it starts increasing linearly!

- I knows that vth is change, then vdsat is also changed.

  But isn't it true that vdsat is definition of vgs-vth?

  Thus I think vdsat equal to vgs-vth at the time of op point.

- In hspice simulation result report file, u can see vdssat and vgs, vds, Vdssat =Vgs-vth , but vds is the real value . For mos working in saturate region, vds must greater than vdssat

- Vdsat=Vgs-Vth is the begin of working in saturate region.

- Most important thing people are forgetting for vds, sat not equal to VGS-VTH is because of two things.

  1. Velocity Saturation.

  2. Hortizontal field due to large VDS in short channel transistors.

  Unfortunately, it is not because of channel length modulation.

  Please refer to any standard circuits book or the BSIM manual.

==

Question:

Relationship between Vdssat and Vov?

As we know, drain-source saturation voltage Vdssat is not the same as the over-drive voltage Vov(=Vgs-Vth). But is there any relationship between them? How can we derive the Vdssat from simulations?

Answer:

- I hope we all agree that Vdsat (min), is the minimum Vds required to keep the transistor in saturation.

  You always want Vds (min)>Vov.

  And you can find Vdsat(min) without simulation, look into 2 stage opamp design example in Allen holberg.

  I am sure you must have observed it that for a simle common source amp having an active current source as load, one cannot give the Vds across it before the simulation. For a tail current source of a diffamp we can surely say what would be its vds as long as input voltage range is know.

- One friend of mine tole me that for long channel MOSFET transistor, Vdssat=Vov, but for short channel MOSFET transistor, this does not hold any more. So it brings forth my question

about the relationship between Vdssat and Vov.

- Vdssat is when it reaches pinch off at the channel. Vgs-Vt is just the potential needed at the gate to inverse the channel under the oxide. This is usually a function of the doping under the oxide and substrate potential.

For short channel, you have a much earlier Vov compared to long channel. This is because the distance between the drain and source is much nearer and thus it creates a much higher electric field. This eletric field causes the velocity to max out at a lower Vds, and that inturn causes pinch off.

- As per your question , VDSsat is drain to source voltage for a MOS transistor. Now that is VDS and not VDSsat as I understand. Vov= Vgs-Vth

As VDSsat(min) is observed lower than Vov (e.g by a 10 to 20 mv) and transistor is shown in saturation. Kindly observe the gds of the transistor for that Vds (<Vov). And compare the gds when it Vds is more than Vov we have designed for. We design for MOS transistors as good current sources.

VDSsat(min) we all find at times less then Vov for which we have designed the transistor cannot use Level 49 equations to do hand calculations.

Short channels do suffer from quick velocity saturation, but for pinch off one still needs Vds>=Vgs-Vth

- We always use Vov to calculate the W/L.

You would observe when you see the DC operating point parameters that VDSsat(min) is very close to the Vov what you have designed for, once the MOST is in saturation.

For long channel, Vov = Vdsat, IV curve looks like sqare law.

If same Vov as above for short channel device, Vdsat<Vov because higher field & IV curve looks like linear.

- Well I guess the need to search for those relation ship is to design your circuit in a more predictable manner.

Because of velocity saturation in submicron devices vdsat<vov, the transistor enters the saturation much earlier.

- I've forgot the formal answer to this question but i know if you wanna receive a complete answer to this, please read the text book "operation and modelling of mos transistor" by Y. Tsividus.

If you just need a 1-st order evaluation, then worry only about vov.

If you're running simulation, please compare vds and vdsat.

If you want to calculate the output resistance, use the reciprocal of gds.

(The latter 2 points were what i learnt from my boss. I don't remember the reasons anymore.)

==

Question:

Vdsat in Spectre?

We can see the parameter Vdsat in Cadence Spectre after we perform the

DC simulation. But I can't figure out what's the physical meaning of it?

I guess the vdsat is the overdrive voltage at first, but it's not exact the same

as Vgs-Vth. The Vdsat is smaller than Vgs-Vth. So anyone has the idea what is Vdsat?

Why not vds>vgs-vth ? How to decide the vdsat?

Answer:

- I dont know, but when I want to see does circuit work well I look at Vdsat. Until now I did not have any problems. I suspect that equation Vgs-Vth is maybe first order approximation?
- I think Vdsat is the voltage at which the carriers travel with saturated velocity not VGT...
- Hi, did you get the Vdsat using Spectre command line?

  In hspice, Vdsat≠Vgs-Vth, too.There is equations for Vdsat in Spice model files.
- For short channel devices,you should make some modification of Vdsat.

  Vdsat=(Vgs-Vt)//(L*Esat)

  Esat is the saturate Electircal field of device, you can get it from model parameter Vsat=u*Esat/2
- Vdsat is the minimum voltage that is required to keep the transistor in saturation.

  In analog design it is a important parameter. Because if the Vds is less then Vdssat then it can cause the transistor to go into triode region, and the linearity and other imp parameters changes.
- Vds = vgs -vth and the conditions what Pixel stated are right to analyse the operating state of Transistor.