VLSI Data Conversion Circuits Prof. Shanthi Pavan Department of Electrical Engineering Indian Institute of Technology, Madras

Lecture - 6 Distortion in a Sampling Switch

So, this is VLSI data conversion circuits, lecture 6. Let me quickly summarize what we were discussing the last time. Around last time we saw that the first non ideality of this of the sampling switch that you can think of is finite on resistance. So, when the switch is closed it does not quite behave like an ideal short circuit there's some resistance. And every physical resistor which is in thermal equilibrium with its surroundings has a noise source associated with it. The spectral density of the noise source turns out to be 4 k T times the value of the resistance.

(Refer Slide Time: 00:50)

And when the switch, when the sampling switch is closed the actual voltage across the capacitor is not simply the input voltage which is been filtered by an R c network, but also the noise of the resistor which also undergoes filtering by the same R c network. And when we calculated the mean square value of this noise voltage across the capacitor we said after a little bit of maths that the mean value of the noise as we expect should be 0. And the mean square value of the noise is the number k T by c volt square and quite surprisingly at first sight it seems as if this mean square noise is independent of the resistor it is surprising, because a the only noise is coming from the resistor.

So, it is a little bit surprising that the mean square noise is independent of the resistor. And only depends on the capacitor after little bit of thought we saw that yes it does make sense. Because even though the noise source strength increases with increasing resistance the bandwidth through which this noise is filtered has reduced. So, these are 2 opposing effects and magically they seem to cancel out resulting in a value which is independent of the resistance. Now, this is somewhat bad news I must say, because now, with our sampling circuit we have definitely picked up the signal that we want to sample. We have also picked up along with it a noise.

A noise a which is 0 mean b which has a mean square value of k T by c. And towards the end of the last class, we discussed that the samples of this noise from instant to instant of the sampled output are uncorrelated. And the reason b that by design 1 would choose this R c network to have a bandwidth which is much higher than the sampling rate correct. If you want to capture the input signal accurately on the capacitor then you must make sure that the resistance of the switch is sufficiently small that the voltage across the capacitor is not an attenuated version of the input. So, by design you would make the bandwidth of this R c filter much higher than the sampling rate which means that the time constant R on time c would be much smaller than.

Sampling time.

The sampling time which basically means that any voltage which was sitting on the capacitor. Before you close the switch will get completely discharged in the next cycle you understand. In other words if the R c time constant is chosen to be much smaller than the sampling time, the capacitor does not remember the previous voltage to a sufficiently large degree. So, pretty much it has forgotten what was there in the previous sample which is equivalent to saying that there is no correlation between successive samples. The sample noise in successive samples if there is no correlation between successive samples. It means that that discrete time process is also a wide process and in the discrete time spectrum that will correspond to a flat noise spectral density.

Now, from this argument we saw that it is not simply enough to have an R on time c time constant which is sufficiently small. As we discussed the other day there are many choices of R on and c which will result in the same time constant. So, we would like to understand if there is a rationale to a specific choice of the sampling capacitor. And the

consideration of noise gives us that answer if you choose the capacitor to be too small what would happen?

Larger variance.

You will have a large variance for the sample noise on the other hand if you choose a small capacitor I am sorry what did I say? If I choose a large capacitor the variance of the noise is small and vice versa. So, we know that after we sample the input on the capacitor it is going to be quantized this sampling completes the sampling in time or quantization of time the next thing is to quantize amplitude. So, now, once we have the sample value it is going to be quantized. So, how do you think? Have you choosed this k T by c, you understand the question? So, let us say whatever voltage is sampled on the capacitor at this point. Let us say I have an ideal quantizer following the capacitor the step size of the ideal quantizer is known because that is I mean that is what you wanted to design in the first place. Now, the question is how will you choose the size of the capacitor? So, what you will do is you will make sure that the RMS quantization noise I mean RMS thermal noise is.

S.

Smaller than the step size of the quantizer, the idea behind this argument is that anyway you are going to lose some information. After you quantize levels you got to make sure that the noise that you add due to sampling is smaller than the step size that you use for quantization all right. And therefore what happens to the; pardon.

It is greater than or

Well, there is there is the question he ask was why do I want to make this choice?. So, what if the RMS quantization noise I am sorry the R m thermal noise is greater than the quantization noise what comments do you have?

Sir, particular thermal.

Let us say for argument sake that the RMS thermal noise is say twice the step size of the quantizer which follows so.

Sir output on the, that quantizer will not be correct as per the.

That is true in either case right whether the.

Sir it is interdependent.

See whether the quantization noise is smaller than thermal noise or not we must accept that there is going to be an error between the input and the sampled output. Whatever we do, we cannot avoid that error correct. Now, the what we are arguing about is to see if it makes sense to make the thermal noise smaller than the step size or larger.

Thermal noise is being.

Thermal noise is closer I mean maximum thermal noise is equal to the step size. You would not be using a bit in information the solution will not be degraded effect of the solution. So, let me rephrase his question. So, what mistake am I making if I say the thermal noise is may be 3 step size. The RMS thermal noise is 3 times the quantization step, what argument do you have to say that I mean from what you are saying it appears as if it is a bad choice correct. The question is why is it a bad choice? I agree that there will be you know more error and so on but.

Sir possibility that it will enter the.

Yes.

Thermal noise the less chances of going into the.

Yes. So, the argument can be rephrased in this way. So, if the RMS thermal noise is much larger than the quantization step. Then I can say that I am not the as per the whole process is concerned there is not much lost if I increase the quantization step is not it because I am adding. So, much error up front that it does not make sense for me to quantize. So, fine anymore, because I have already added. So, much error in the first place. So, I can use a quantizer with a much coarser step without really effecting the entire system, you understand, does it make sense?

There are 2 errors which you are adding 1 is the k T by c noise that you add when you sample and 2 the quantization noise due to the quantization that follows. Now, the question is why should I choose one to be you know much smaller than the other. So, if I chose the thermal noise to be larger than my quantization step. Then somebody could argue that why do I then need to spend such a lot of effort in trying to build a quantizer with a small step presumably. It is much easier to build a quantizer with a large quantization step than with a small quantization step, does this make sense? So, if you want to resolve to a much more accurate value; that means, that the quantization step has to be very small, but if you take the argument that I am going to have.

So, much thermal noise right up front that it is RMS value is much larger than my quantization step size. Then as a quantizer designer I would say then why waste time and effort trying to design a very fine quantizer. Anyway you have made such a big error up front. So, I will make my job easy and make a quantizer with a much smaller step size. A much larger step size correct you understand. So, if you want a final resolution for the entire converter for the entire signal path both sampling in time as well as in amplitude. Then you got to make sure that the error you pickup. Finally, from end to end is bounded right by some value what we assume is that the sampling process results in a much smaller error than the quantization process. So, then you can say that the entire resolution of my chain is simply that of the quantizer, is that clear? yes.

Sir, the RMS that you had considered or some peak to peak.

Well, so, now, now the question of whether we should consider RMS or whether we consider peak to peak is an open question? But 1 I mean some measure of the thermal noise must be smaller than the step size. When we smaller than the much smaller than step size. And again now the question is what is much smaller right a common value to choose is may be a quarter of the step size of the quantizer you understand? So, all right now what implication does this have on design as far as high resolution a to d conversion is concerned? So, if you have a limited signal swing, and if you want to resolve the signal to a very fine value in other words the resolution of your entire a to d conversion process is high. Then it must follow that the capacitance that you use for sampling is.

Very large.

Sufficiently large. So, that the RMS thermal noise introduced due to sampling becomes a small fraction of the step size. The step size you recall is now much smaller, because you are targeting a very high resolution, you understand?

(Refer Slide Time: 14:55)

So, if for example, the peak signal amplitude is a then one can quantify the signal to noise ratio after the sample and hold as ratio of the signal power to the.

Noise.

Noise power the signal power for a sinusoidal input will be how much?

A square by 2.

A square by 2 the noise power is k T by c. So, the signal to noise ratio after sampling is a square time c by 2 k T and this better we much larger than the required. The required signal to noise ratio of the whole chain which consist of the sampling operation plus the quantization operation does it make sense. So, the next thing to do is to try and figure out how one might be able to realize this using devices that we know. And as you are all aware the mosfet is an ideal candidate for this kind of work, we know from earlier courses that a mos transistor.

(Refer Slide Time: 16:37)

So, the mos transistor is an ideal candidate for used as a switch. Let me quickly refresh your memory an n mos transistor which was shown here has 3 terminals the drain the gate and the source and the body. The body is usually not accessible in many technologies and it is a condemned to be connected to ground. So, typically as I was mentioning the body is connected to ground. And one has no access to the body as such technologies where this is possible for the n mos transistors are what are called n, well c mos processes. And what reason do you think you would want to operate the transistor in? If it wants, if you want to make it behave like a switch.

Sir, k T value equal to linear.

You want to operate the transistor in the triode or linear reason when you want to make a closed switch. And you want it to be off when you want to make an open switch.

(Refer Slide Time: 18:39)

And let me quickly put down expressions for the resistance of the switch etcetera, when its operating as a closed switch. So, what is the limits for triode operation?.You got to make sure that V g s.

Greater than.

Is greater than.

V T.

V T all right and V d s is less than or equal to V g s minus V T and the threshold voltage itself is not really a constant it depends on the.

Source to substrate.

Source to substrate bias and V T is therefore, given by some relationship of the form V T naught plus some gamma times square root of V s b plus.

2 phi.

2 phi naught minus square root of 2 phi naught. All this comes from the theory of mos transistors and is not, we would not go into these details. But as far as a designing a sample and hold is concerned we need to know these expressions. And in the triode region the current in the mosfet is given by mu n c ox w by l times V g s minus V T times V d s minus alpha V d s square by 2 where all terms have their usual meanings. And this alpha accounts for the variation of the threshold of the transistor along the channel length nd is typically about 1.1 or 1.2. In most cases that we will deal with we will be operating with such a small V G S across the switch with a such with such a small V d s across the switch that.

(Refer Slide Time: 20:45)

One can actually neglect this term in which case the resistance of the switch is approximately given by 1 by mu n c ox w by l times V g s minus V t.

If alpha is.

Pardon.

In which we see alpha is 1.

Yes

No, no actually the comment made was now only we see half V d s square in the equations. And where as this alpha coming from it turns out that half V d s square is not entirely correct right. There is the correct equation is to have those of you who have taken done a mos modeling class or a device modeling class will know this that the standard way of deriving the I from the V g s. And V d s is to assume that the channel charge is of the form V g s minus V T. And that V T is assumed to be independent of the location along the channel in practice it turns out. I mean physically you can see that as you go along from the source to the drain that the depletion division must actually the depletion region thickness must actually increase, because the source substrate potential is smaller than the drain substrate potential. So, threshold actually increases as you keep going from the source to drain. So, in that this alpha models that effect anyway do not worry about it if you taken a mos class. And you must bear in mind that V T is a function of the source substrate biased voltage through the body effect. So, now, let us come to our first version of the.

(Refer Slide Time: 23:01)

Sample and hold circuit all that we have done is replaced the switch with something which opens and closes based on electronic control. And that is the gate potential further we assume that the whole circuit. And all circuits that we deal with operate between 0 and V d d which means that all our signals will nominally be at V d d by 2 and will jump up and down around V d d by 2. So, let me call that voltage V c m is the V d d by 2. So, this signal that we wish to sample will be of the form V c m plus say a sin omega T. Now, when the switch is on V g must be connected to the when you are tracking the transistor must be.

On.

On right. So, when it is tracking you want to make sure that the resistance of the switch is the.

Small.

Smallest possible. So, that you have high tracking bandwidth. So, if you had a fixed transistor size what you would do? Would be to connect the gate to the.

Highest possible

Highest possible potential and the highest possible potential that we can think of is.

V d d

V d d. So, V g must be connected to V d d during the track phase and V g must be connected to ground during the hold phase does it make sense? Now; that means, that V g of T must look like this. This is V d d and this is 0 and this corresponds to track and this corresponds to hold Now, let us see what nonidealities there are in this circuit if everything works fine. We just move on we are done right if things do not work properly we need to see what the problems are and fix those problems. So, 1 thing that 1 should notice is that the mos transistors resistance depends on V g s minus V t. So, if you plot the wave form at the input, what does this wave form look like? It is simply a sinusoid hovering around this is time this is V c m; this is V c m plus a right Now, can you comment on the gate source voltage of the transistor. So, the gate during the track I mean of course, the gate potential is varying right during the track phase can you comment on the gate source voltage..

What is the gate source voltage?

In some system V d d.

V g the gate voltage is simply V d d during the track phase.

(Refer Slide Time: 28:50)

What about the source voltage?

V c m plus.

It is V c m.

Plus.

Plus whatever input there is let me call this a sin omega T correct.

Sir.

So, what is V g s?

The sampled numerous in the source strain.

Pardon.

How sample or output.

During the track phase, what is happening is that let me since there seems to be is some confusion limit during the track phase. The gate is connected to V d d the input is connected as such V c m plus a sin omega T and the capacitors here. So, if the switch resistance is chosen to be sufficiently small which is what one would do in practice then they drop across the transistor should be.

Small.

Very small compared to the drop across the capacitor correct. So, the source gate voltage or the gate source voltage of the transistor is nothing but V d d minus V c m plus a sin omega T which is V d d by 2 minus a sin omega T. What do we notice do you think is there a problem time constant? Time constant is a function of the input.

Pardon.

Time constant is a function of input.

Is a function of input what?

Amplitude.

So, basically I mean one thing you notice is that the on resistance of the switch is proportional to inversely proportional to V g s minus V T. As the input keeps changing the gate potential is fixed in the track phase. But the source potential is varying that is basically the input you want to sample if the input becomes too high. What is happening to the 8 source voltage?

It reduces.

It is reducing right, if the gate source voltage reduces. What happens to this quantity V g s minus V t?

Reduces.

Reduces, if this reduces what happens to the resistance of the switch?

Increases.

Increases right and on the other hand when the input is in the negative half of the cycle the gate source voltage is increasing which means that the R c time constant is decreasing. And one extreme case is what he pointed out which is if the input voltage becomes too high you stand the danger of?

Turning off.

Turning off the transistor during some part of the cycle correct because if this potential becomes so high that this drop becomes less than V T. Then the transistor will be turned off you understand. So, this is a problem because the resistance of the switch is now a function of input amplitude. So, is this a linear phenomenon or a non-linear phenomenon.

Nonlinear.

It is a non-linear phenomenon which means that it can result in what constant harmonics harmonic?

(Refer Slide Time: 33:49)

It can result in harmonics; it may result in distortion you understand? So, let us try and understand in a more quantitative manner what happens. So, ((no audio 33:43 to 34:31)). So, in general we see that this resistance is non-linear intuitively?

(Refer Slide Time: 34:52)

Let us try and quickly sketch what 1 one must expect for the output to look like if the input was like this. And the resistance of the switch was larger when the input went up and smaller when the input went down, what do you think? You should expect if the resistance becomes larger when the input goes up. What does that mean for the bandwidth of the R c network?

It reduces.

Reduces if the bandwidth reduces what do you think will happen to the delay between the input sinusoid and the output sinusoid?

Increases.

The phase shift of the delay increases where as when the resistance decreases from the nominal value.

Right.

The phase difference between the output sinusoid and the input sinusoid.

Reduce.

Must reduce. So, what you can expect is some waveform with a bad rendering, but I hope you get the idea.

Sir certain case.

Yes.

It is good natural good know sir.

Well, I mean I do not know about good or bad, but one thing you must be make this is clearly not a linear phenomena that is what I am trying say. I mean this is mostly exaggerated picture for you to be able to see and understand. But this is what? This is the kind of wave form that you must expect as you can see the phase difference between the input and output is much higher for larger values of amplitude. And smaller when the input goes in to the negative half of the cycle in general you can kind of model in order to get an idea of what this does to the distortion. One can model the resistance of the switch to depend on the input voltage using a polynomial. So, in other words let us say the resistance depends on the input voltage. So, let us say this is V c m plus a sin omega T under nominal conditions when A is 0 the voltage drop across the switch what do you think would be the voltage drop across the switch?

We assume it is.

Pardon.

We assume it is small.

We assume it is small, but when let us see. So, if the voltage drop across the switch is very small then what is the voltage drop across the capacitor approximately?

Input.

Say roughly the input. So, the output voltage is approximately V c m plus a sin omega T which means what is the current through the capacitor approximately?

The V c V g s.

It is.

V s approximation.

 C a.

Omega cos.

Cos omega T, this is the displacement current through the capacitor. Now, why is there a drop across the switch? And it is there is a drop across the switch, because the switch has got finite resistance. And there is some current flowing through the switch and the current flowing through the switch is approximately.

C A omega.

This current is approximately C A omega cos omega T you understand. So, if the switch was perfectly linear the resistance of the switch would be constant. In other words the voltage drop across the switch would be what?

Sir remaining.

Current through the switch is C C omega cos omega T the resistance of the switch if it was constant would be some R on. So, the voltage drop across the switch would be.

R d omega.

R on into C A omega cos omega T. This is an approximate relationship given that we have assumed that the voltage drop across the switch is small in relation to the amplitude of the amplitude of the voltage wave form across the capacitor does it make sense. You will get the same thing if you use Laplace transforms and use the approximation that omega is much smaller than 1 by R c. I mean, what does it mean physically to have a large tracking bandwidth. It means that the voltage drop across the switch is much smaller compared to the input which means that the output amplitude will roughly be the same as the input amplitude is this clear? Now, unfortunately this is this is what you would have for a linear switch.

Now, unfortunately the resistance of the switch is no longer linear and it depends on the input amplitude. So, let me say the resistance of the switch is of the form is a function of input amplitude. And is of the form some R on which is a constant plus something which depends on a plus something which depends on a square and so on. I mean its non-linear we know that for instance if the input amplitude is positive or rather the input is positive. Then let me not confuse you by calling this a let me call this V in. So, that and if V in is chosen to be a sinusoid of a amplitude a then all these relationships hold.

(Refer Slide Time: 42:36)

So, instead of calling this A I will call the resistance of the switch which is now dependent on V in And why does this resistance being dependent on V in make sense is because the mass resistance is being modulated by the input voltage. So, it is R on plus R 1 times V in plus R 2 times V in square plus higher order terms. Hopefully the first 2 terms are good enough to modulate reasonably accurately or at least for us to get an idea of what effect the nonlinearity of the switch has on the wave forms. So, if the resistance is expressed in this form do you think the, for the specific case we are talking about do you think R 1 is positive or negative.

Positive.

So, if R 1 if V in increases if V in is positive the resistance.

Increases.

Increases. So, R 1 is.

Positive you understand. So, depending on the nature of the device we can go and evaluate what this R 1 and R 2 will be you understand? In fact, with the mos transistor now this is the resistance of the switch V g s can be expressed in terms of V d d minus V. in V d d I mean as V d d minus some V c m.

Minus V c m.

Minus V in all the stuff V T is also dependent on.

V s b.

V s b which is V c m plus V in correct. So, both V g s and V T depend on the input through some complicated expressions. But the bottom line is that the resistance depends on the input and can be in a model 2 a first order like this. Now, if this is the resistance what do you think the drop will be across the resistance this with the, into current. The voltage drop across the switch is simply a current flowing through the switch times its resistance. The resistance now happens to depend on V in correct. So, and the current is what the sampling capacitance d V in by d T times R on plus R 1 times V in plus R 2 times V in square. And this is an approximate relationship why is this an approximate relationship where is approximation here, because while calculating current, we assumed that the.

So, there are the, this is an approximate relationship, because we have used c d V in by d T to find the current to the capacitor where as the true current through the capacitor is what C d g out. The o the current through the capacitor is c d V out by d T right while you are saying is that if the drop across the switch is very small. Then V out and V in are about the same which means that the current through the capacitor can be approximated by c d V in by d T does this make sense? Now, c d V in in d T is the current. And this is the resistance the drop is simply the product of the 2. Now, let us try and evaluate this for a special case of a sinusoidal input. So, if V in is say a sin omega T or omega in times T. Then the voltage drop across the switch becomes c a omega in cos omega in times T times R on plus R 1 a sin omega in times T plus R 2 a square sin square omega in times T does it make sense. And what is this c a omega in times R on cos omega in times T plus c a square R 1 omega in times cos.

(Refer Slide Time: 48:11)

Omega in times T times sin omega in times T plus c a cube R 2 omega in cos omega in times T times sin square omega in times T. So, what do you think these terms are?

R 1.

So, 1 can replace 2.

Into.

Times cos omega in T sin omega in T sorry 1 can replace the product sin and cos by.

Sin 2 omega.

Sin 2 omega in times T and what about this character?

1 minus 2.

1 minus 2 sin square omega in times T is cos of 2 omega in times t. So, which means, that this now can be replaced by 1 minus.

Cos.

Cost 2 omega in times T and this in can be further simplified into c a cube R 2 omega in by 2 times cos omega in times T minus c a cube R 2 omega in by 4 times.

Cos 3 omega in. Cos 3 omega in t. Plus cos omega. Plus. Cos omega. Cos.

Omega in t.

Omega in T. Now, before we you know get some more intuition on this one thing I would like to point out is that if you want to build a fairly decent sample and hold. Can you comment on the various terms here the relative magnitudes of R 1 V in and R 2 V in square versus R on.

.What are dimensions of R 2?

Ohm per.

Ohm per volt square and R 1 ohm per volt. So, if you want to make a fairly decent sample and hold. Iin other words a fairly linear 1 what will you try and attempt by design.

To reduce.

You will try and make this term and this term much small compared to what?

R on.

R on right. So, in other words R on you try and make it larger than or you choose and try and make the switch such that R 1 V in and R 2 V in square. The drops associated with those terms the non-linear terms is much smaller than the drop associated with the linear term. Now, therefore, it if you kind of try and use that intuition here the drop across the switch can be simplified to be C A omega in times R on cos omega in T. And what else there are other terms which have cos omega in T 1 is c a cube R 2 omega in by 2. And the other term corresponding to omega in is c a cube R 2 omega in by 4r. So, by the discussion we just had regarding if you had a fairly linear switch which to begin with then these terms will be very small in relation to the term contributed by.

R on.

.

R on. So, you can kind of neglect those in comparison to the term contributed by R on does it make sense. So, this is which means that we can now get rid of these 2 characters we are left with the component at 2 omega in and 3 omega in and they are somewhere down the line we have to design in such a way that they are not yet.

(Refer Slide Time: 54:06)

Because they are term fine that is the idea right eventually if you want to make a linear switch you want to make sure that these R 1 R 2 become accordingly. I mean see when we are talking about the first macro model of the switch we just assumed finite on resistance the only problem that that causes is tracking bandwidth. And thermal noise is of course, a fundamental thing there is nothing you can do about it. However, now if the switch is non-linear in other words the resistance of the switch depends on the input. Apart from tracking bandwidth issues the drop across the switch also has components which are harmonics of the input frequency right. So, this represents the drop across the switch let us try and stare at this expression. And see if this makes why this makes sense? First why do you think there is a term which is proportional to omega in why do you think that makes sense? The drop across the switch is proportional to the current the current is given by.

 C_{\cdot}

C d V d T right. So, as the input frequency increases the current through the capacitor increases causing a larger current in the switch therefore, increasing the drop across the switch. Now, what does it make sense that the second harmonic current is proportional to the input amplitude square where is the second harmonic term coming from?

V in square onto V in square.

It is coming from.

V in square.

The V in square and where is this V in square coming from?

R 1 V in resistance.

The second harmonic is coming from R 1 V in multiplying with d V and d T you understand? So, the second harmonic is coming, because a the current has a component a the resistance is varying and the amount of variation also depends on A. So, the drop is this multiplied by that. So, that is why you get the A square similarly, the third harmonic is coming from the R 2 V square term correct.

Sir, then it is a harmonic 2 omega in.

Yes.

Is harmonic frequencies.

What harmonic means frequency?

Sir.

No I am saying the amplitude of the second harmonic is proportional to a square is this clear? And why it make sense is because the second harmonic is coming from this component of the resistance multiplying with the current as the input amplitude increases this goes up by a factor of A. So, does the change in the resistance? So, when you multiply the 2 you get a second harmonic component which is proportional to a square similarly the third harmonic component is is proportional to.

A cube.

A cube and the way you would have designed the track and hold. You would have tried to make this I mean this is a linear term that is the voltage across the fixed part of the resistance and that is a linear term it does not cause harmonics. So, as we expect because of the nonlinearity of the switch there are harmonics which is the characteristic of a nonlinear system you put in a sinusoidal frequency of some omega in. And there are voltages and currents in the network which are multiples of which are at multiples of omega in does it make sense.