DVCon India 2024

TITLE OF PAPER	**Power Probe: Addressing Power Noise Signal Integrity Challenges for Wide IO HBM Memories Through Advanced Verification Approach**
AUTHOR 1	**Name: Akshobhya B** **Organization:** Samsung Semiconductor India Research(SSIR) **Job Title: Senior Engineer** **Email ID: akshobhya.b @samsung.com** **Mobile no: +919901620581**
AUTHOR 2	**Name: Bhargava Krishna Venigalla** **Organization:** Samsung Semiconductor India Research(SSIR) **Job Title: Associate Staff Engineer** **Email ID: b.venigalla@samsung.com** **Mobile no: +917075494775**
AUTHOR 3	**Name: Giridhar Rangarajan** **Organization:** Samsung Semiconductor India Research(SSIR) **Job Title: Senior Engineer** **Email ID: g.rangarajan@samsung.com** **Mobile no: +916362940060**
AUTHOR 4	**Name: Gowdra Bomanna Chethan** **Organization:** Samsung Semiconductor India Research(SSIR) **Job Title: Associate director** **Email ID: gb.chethan@samsung.com** **Mobile no: +919538367653**
AUTHOR 5	**Name: Anil Deshpande** **Organization:** Samsung Semiconductor India Research(SSIR) **Job Title: Director** **Email ID: anil.pande@samsung.com** **Mobile no: +919591108099**

Abstract:

We are currently existing in an era where terms like Artificial Intelligence(AI) and Machine Learning(ML) have become common household jargon. To cater for the ever-growing demand for processing large amounts of data, High Bandwidth Memory(HBM) was introduced as a solution. HBM addresses the demands like High Bandwidth with low latency memory accesses by providing a wider data bus with upto

64 bits of DQ, this means that the amount of data being processed per clock cycles is much more resulting in higher speed and efficiency. Since HBM memory is a 2.5D memory consisting of memory dies interconnected using through-silicon vias (TSVs), it offers relatively less latency than other traditional interconnection methods and higher packing density. However, this results in an increased amount of IOs opening up a larger margin for external as well as internal factors to affect its normal operation. Here, we are discussing the causes and resultant effects of one such prominent factor, **power noise**. We will be modelling its effects in the digital domain and simulate the results.

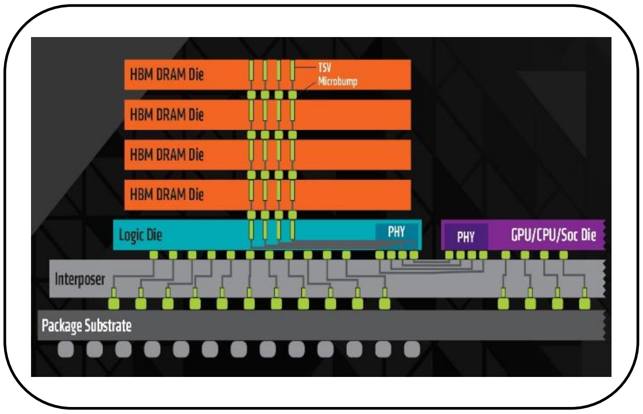


Figure 1: HBM memory architecture

Power noise, in of HBM (High Bandwidth Memory), refers to fluctuations or disturbances in the power supply voltage experienced by the memory devices, PHY (Physical Layer), and associated components. These fluctuations can occur due to various factors such as rapid changes in current demand, impedance mismatches, parasitic elements in the power delivery network, and electromagnetic interference.

The ***effects*** of power noise in HBM memory can be significant and can impact the memory system in several ways:

1. **Bit Errors and Data Corruption**: Power noise can induce fluctuations in the voltage levels of memory cells, leading to bit errors during read or write operations. These errors can corrupt stored data, resulting in data integrity issues. This also leads to increased bit error rate(BER).

2. **Timing Violations**: Power noise can disrupt the timing margins of memory operations, leading to violations of setup and hold times. When the timing constraints of memory interfaces are violated, it can lead to data setup failures or hold failures, resulting in data corruption or loss.

3. **Reduced Operating Margin**: Power noise can reduce the operating margins of the memory devices, causing them to operate closer to their specified limits. This reduction in margin can make the memory system more susceptible to variations in operating conditions, such as temperature fluctuations or voltage droops, potentially compromising its stability and reliability.

4. **Interference with ECC** (Error Correction Code) **and SEV** (Severity): HBM memory often employs ECC techniques to detect, correct memory errors and Severity pins to indicate severity of errors. However, power noise can interfere with ECC/SEV mechanisms by corrupting parity bits or introducing errors in the calculation process, reducing the effectiveness of error correction and potentially leading to undetected errors.

Power noise in HBM can arise from various ***sources***, including:

1. **Switching Activity**: Rapid changes in current demand due to the switching of data signals and control signals within the memory devices and associated circuitry can generate power noise. The high-speed nature of HBM operation exacerbates this effect.

2. **Impedance Mismatch**: Mismatches in impedance between the power supply and the load can lead to reflections and voltage spikes, contributing to power noise. This can occur due to variations in impedance along the power distribution network or from differences in load characteristics.

3. **Inductance in Power Delivery Network**: Inductance in the power delivery network, such as inductance in power traces and package leads, can cause voltage drops and ringing during rapid changes in current demand, leading to power noise.

4. **Crosstalk**: Crosstalk from adjacent signal lines or power rails can induce noise in the power supply, affecting the stability of HBM circuits. Crosstalk can occur due to coupling between closely spaced signal lines or due to shared ground and power planes.

5. **External Interference**: External sources of electromagnetic interference (EMI) or radiofrequency interference (RFI) can couple into the power delivery network, introducing additional noise and exacerbating power noise issues in HBM.

6. **Affects DLL locking**: Power noise in HBM can have detrimental effects on the performance of a Delay-Locked Loop(DLL) by introducing jitter into the clock signal used by the DLL. Jitter refers to the variation in the timing of clock edges, which can affect the DLL's ability to accurately lock onto the desired phase. Power noise-induced jitter can lead to increased uncertainty in the DLL's output timing and degrade its performance.

Application:

Modelling of Power noise: Gaussian White Noise(GWN) is a type of random signal characterized by a Gaussian distribution with a mean of zero and constant power spectral density across all frequencies.

Mathematically, it is represented as:

$N(t)=A\cdot N(\mu\ ,\sigma^2)$

Where N(t) is the GWN signal at time t, A is the amplitude, and

$N(\mu, \sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2.

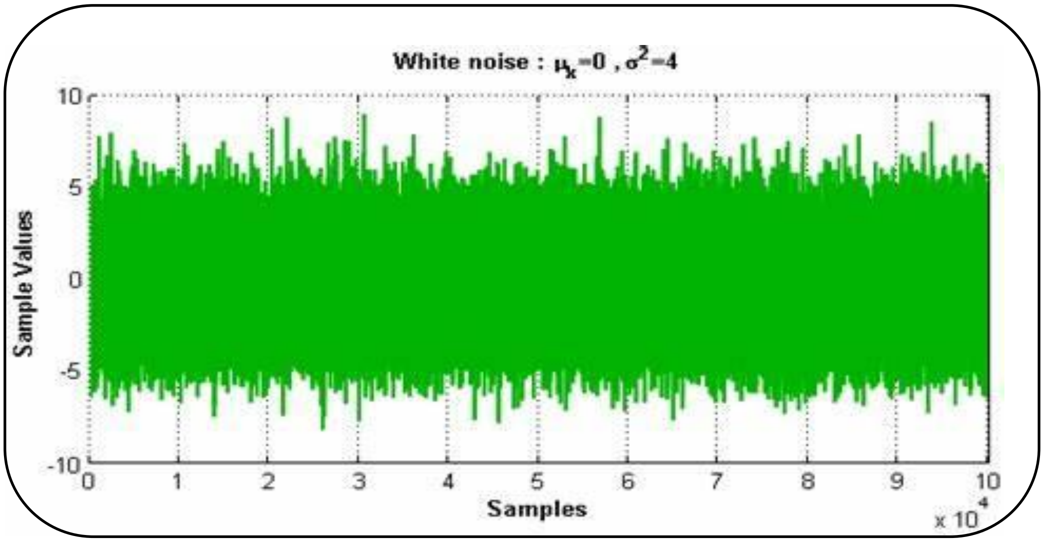


Figure 2: Gaussian White noise simulation

In digital systems, GWN is often generated using pseudo-random number generators (PRNGs) with a Gaussian distribution. One common method for modelling GWN in System Verilog is using the $dist_uniform function to generate uniformly distributed random numbers and then applying a transformation to obtain a Gaussian distribution.

This approach of modelling power noise is dependent on inputs like the coefficient for Gaussian equation that we need to receive from analog team. Additionally, we also noted that a Gaussian approach is slightly inaccurate as there is a lot of randomness.

So, another methodology was proposed and implemented for the same problem. In order to find a correlation between change in VDD (supply voltage) and the amount of distortion caused, simulations of a particular sequence is performed. The sequence is executed as follows –

1) Fix a value of VDD
2) Consider 4 different variations on sum of damages applied to the right and left edge of the data(DQ) signal
3) Run centring training and find out offset value w.r.t sampling/strobe signal
4) Based on tolerance of the offset value, we quantify the result as 1 for a pass and 0 for a fail

Once this data is available, the approach to correlate the two variables is by plotting these values on a schmoo plot.

For this, X-Axis is taken as the VDD, Y-Axis is taken as sum of left and right damage. On this, we plot the results of the centring training as we discussed earlier. A pass (1) is indicated using a green dot and a fail

(0) is indicated using a red dot. Running a linear regression through the green region will give us an almost accurate correlation equation.

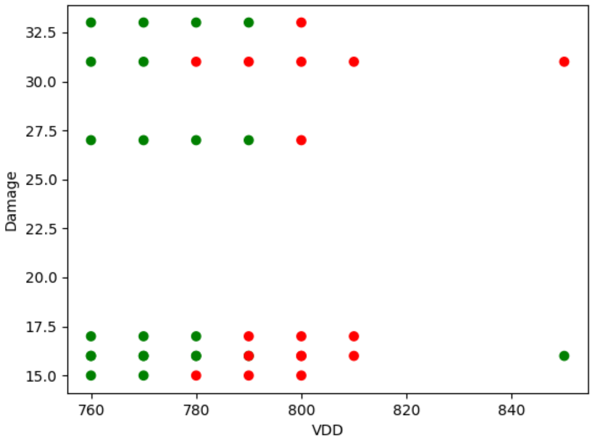


Figure 3: VDD vs Damage Schmoo plot

This is an ongoing effort and a lot more data is needed for conclusive results. Another enhancement would be changing the X-axis from VDD to delta VDD (change in VDD).

The legacy approach being followed is Power noise generation using Sinusoidal equation:

vdd_noise_mv = sin (2*pi*noise_resolution)

noised_vdd = vdd_dc + vdd_noise_mv, is supplied to DUT via PDN UVC to check if the DLL functionality diverges from the ideal behaviour (Figure 10).

Power noise in HBM manifests as transient fluctuations in the voltage and current levels of the power supply network. These fluctuations occur in response to rapid changes in current draw during switching activities, such as read and write operations, as well as high-speed data transfers. By generating power noise using GWN, it is possible to simulate the impact of random voltage fluctuations on the performance and reliability of HBM devices, aiding in the design and optimization of power delivery networks and noise mitigation strategies.

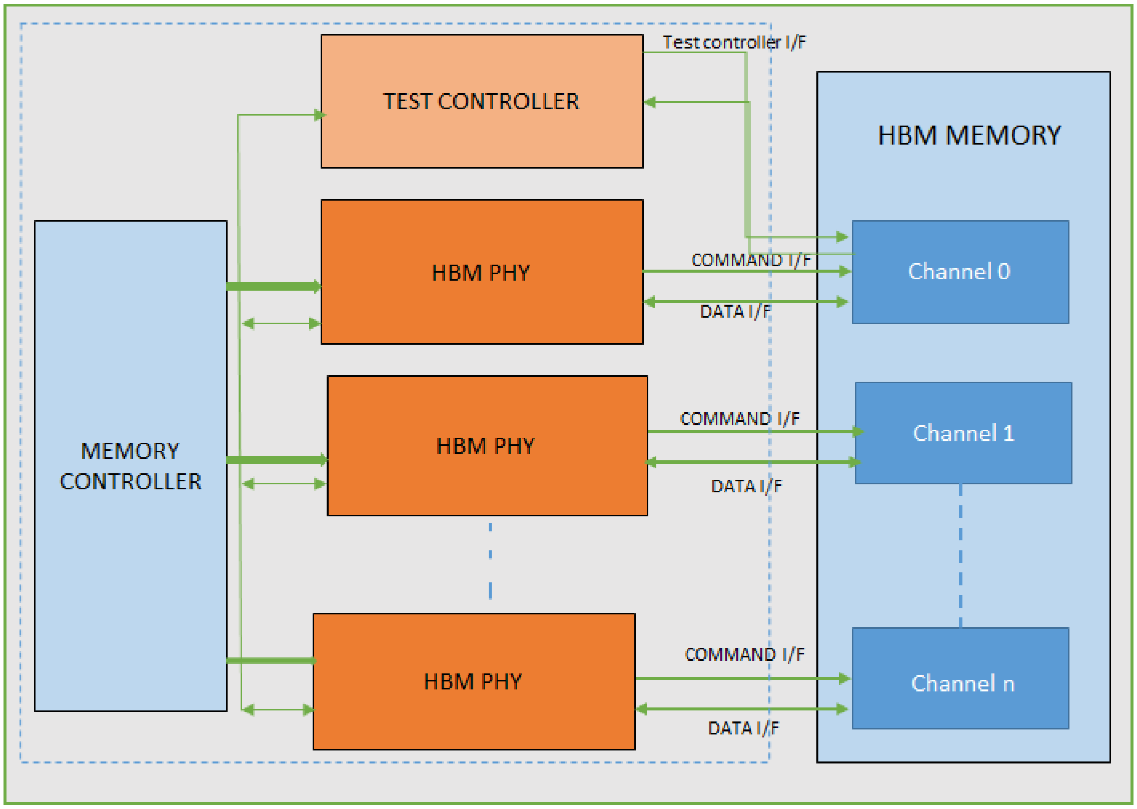


Figure 4: Memory sub-system with multiple PHYs

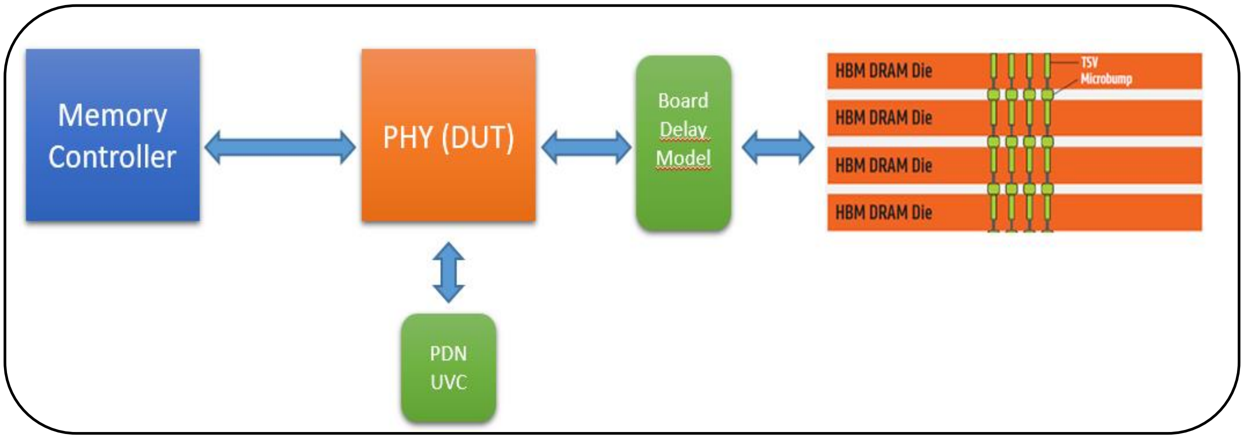


Figure 5: Memory sub-system TB Block Diagram

In the above Figure 5, PDN UVC is used to generate GWN using the general mathematical equation of the Gaussian distribution. This noise could influence the HBM PHY mainly in 2 ways.

1. It affects the Delay Locked Loop by changing the fine delay and hence it might induce jitter in the clock. Jitter effects due to power noise in HBM (High Bandwidth Memory) can cause data corruption and signal integrity issues.
2. At the interface level between PHY and Memory, noise can introduce distortion in the data bus and cause issues like reduced valid window margin and timing violations. Power noise can also lead to fluctuations in the timing of data signals, resulting in increased jitter. This can be particularly problematic in high-speed memory systems where timing is critical. We have used ***Board Delay Model***(BDM) UVC which will add delay/distortion as per the value returned by the Gaussian noise equation. BDM is used to induce jitters on data signals.

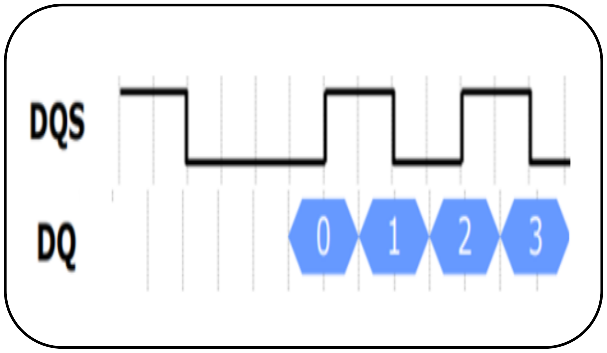


Figure 6: Data Path without distortion

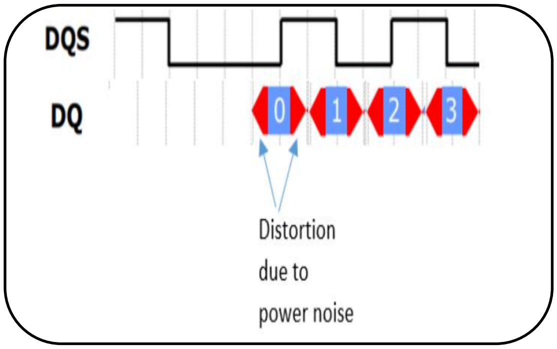


Figure 7: Data Path with distortion

We have developed TB using UVM architecture. The PDN & BDM blocks are developed as UVC components. For validating the Power noise feature with HBM PHY & Memory, we have added checkers to make ensure that there is no data mismatch & ensure that trainings happen properly for capturing optimal data eye.

Further expanding on the TB and its components –

Test: The test is at the top of the hierarchy that initiates the environment component construction. It is also responsible for the TB configuration and stimulus generation process.

ENV: This component encapsulates the entire verification environment, including agents, scoreboards, and other elements necessary for the test. It provides a structured way to organize the TB components.

HBM PHY: HBM PHY is the Design Under Test (DUT). HBM PHY consists of several key components that manage the physical signaling, timing, and data integrity for high-speed memory interfaces.

Virtual sequencer: Virtual sequencer is a special type of sequencer that is used to coordinate the activities of multiple sequencers across different agents. It is useful in complex verification environments where multiple interfaces or protocols need to be verified simultaneously, and their interactions have to be synchronized.

Board Delay Model (BDM): This is a UVC responsible for generating distortion and skew between HBM PHY (DUT) and HBM memory

Power Distribution Network (PDN): This is a UVC responsible for generating the VDD, Finedelay of the DLL and power noise.

APB Agent: APB agent is a reusable, configurable, and self-contained Verification IP designed to verify the APB interface of the HBM PHY. It encapsulates all the necessary components to generate, drive, and monitor APB transactions, ensuring the DUT complies with the APB protocol specifications.

Scoreboard: Scoreboard collects and compares expected results with actual results from the memory model and Memory controller VIP.

Checkers: During initialization of PHY, due to power noise, command and data signals might not be center aligned with clock. PHY trains data and command path in such a way that they are center aligned with clock and we have centering checkers for data and command signals which checks whether data/command is center aligned with clock or not to capture maximum data eye.

Test Controller: It provides the interface connection between host and HBM DRAM, and enables test instruction commands to execute across channels.

Functional Coverage: Functional coverage will measure what features of the design have been exercised by simulation testing .We will be sampling variables in the testbench as per Functional V-Plan to analyse if they have reached specified set of valid values. This verification via coverage approach includes covering the following individually -

1) Maximum and minimum data rates
2) Write, read and parity bit latencies, which are constrained, based on the operating frequency and in turn, data rate of the sub-system.
3) Voltage and frequency range variations

These individual parameters to be covered are called coverage bins and each value that is set to be covered is known as a cover point, further in order to ensure that the particular scenario will performed with all possible legal combinations of these coverage bins, we create a matrix crossing the individual coverage bins combinations. This is also known as a cross bin. Through this, we can also ensure that corner case combinations are also covered.

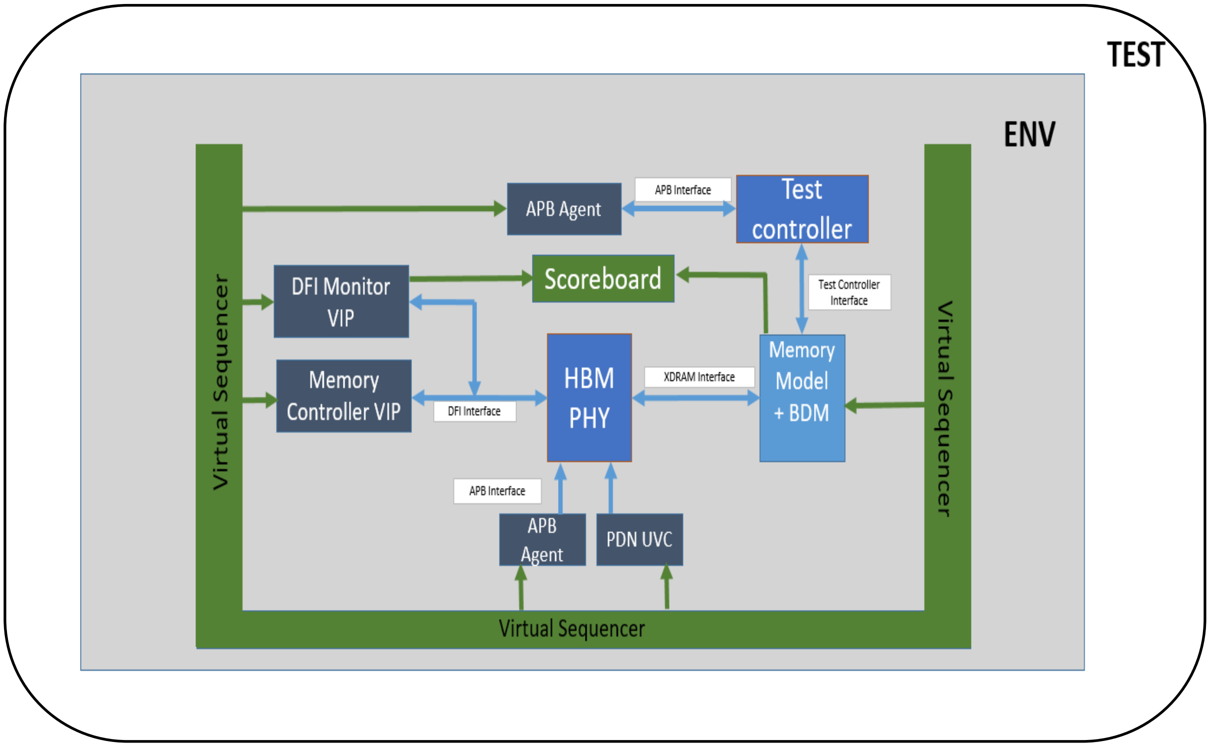


Figure 8: UVM TB architecture

Results & Conclusion:

In this paper, we have discussed about verification of power noise in high performance, low latency and power consumption of HBM memory subsystem at digital domain. By modelling power noise in UVM TB, we were able to mimic real time behaviour effects, which will increase the confidence of Design by catching the issues at very early stages of design cycle.

In the below waveform snapshots, it is shown how power noise is modelled in digital domain.

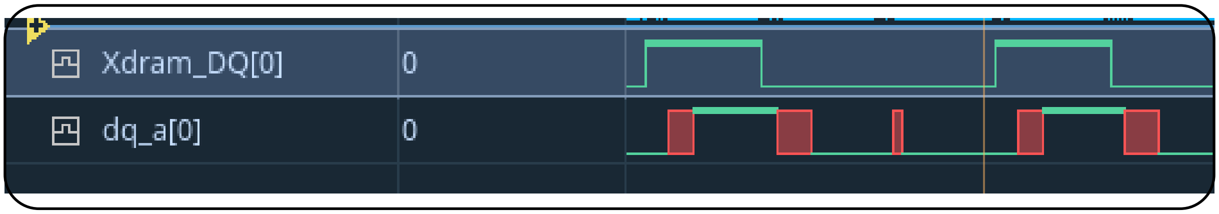
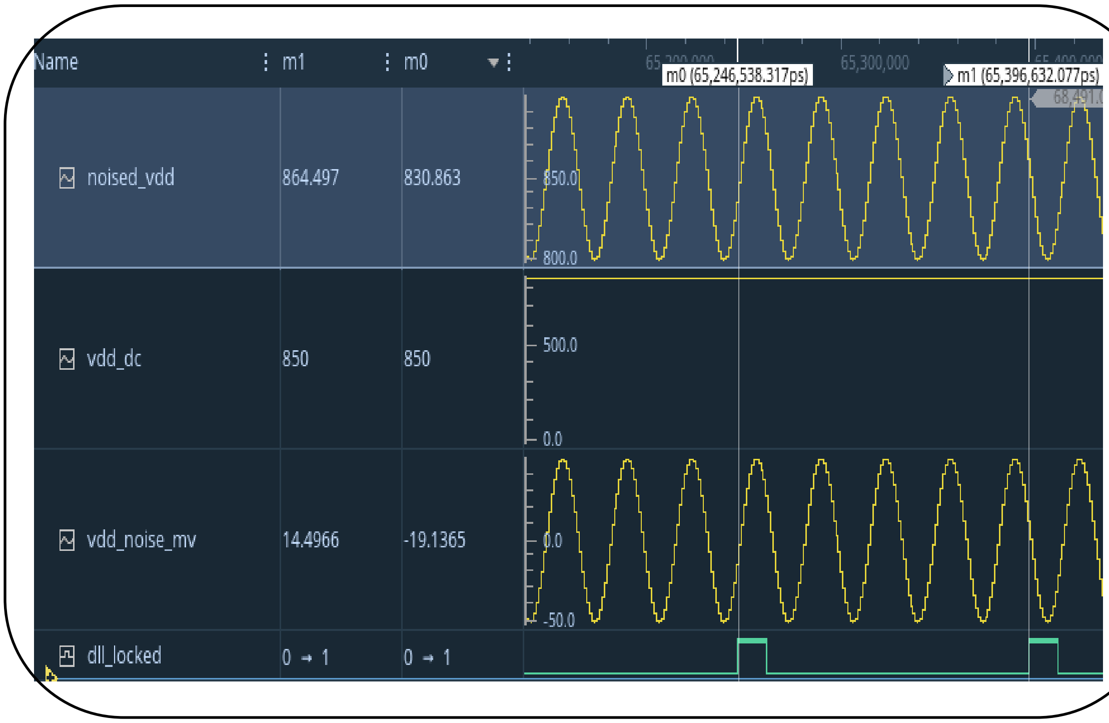


Figure 9: Modelling of Power noise for data interface

In Figure 9, XDRAM_DQ is output from PHY to BDM ; dq_a is the output from BDM to memory dq_a signal is generated by adding effects of power noise. This is done prior to performing Read DQ calibration as we can train the PHY for adverse conditions which could arise due to the discussed silicon effect, power noise. In this scenario, we expect power noise to manifest itself as distortion in data (DQ) channel.

The other ways in which its effects can be seen are fluctuation in voltage (sinusoidal in this case) and an increase in its amplitude, which would in turn result in rising operating temperatures of the sub-system. Temperature, being an important parameter to maintain ensure optimal functionality.

Change in VT (voltage and temperature) parameters, affect the DLL and change the value of finestep delay. Finestep delay is the amount of skew that can be added at once to DQ bus, PHY will keep adding this amount of skew until centering between DQ and sampling/strobe signal is achieved. A higher voltage results in a smaller value of finestep delay. This means more iterations are needed to ensure centering and eventually for a higher value of voltage, we may not be able to achieve centering, as the value of finestep delay would be a very small value. With a fixed number of delay cells, the skew between DQ and strobe signal would exceed the maximum skew we can apply through the DLL. This would affect data integrity, as we would have unreliable sampling.



Name	m1	m0
noised_vdd	864.497	830.863
vdd_dc	850	850
vdd_noise_mv	14.4966	-19.1365
dll_locked	0 → 1	0 → 1

Figure 10: Modelling of power noise for DLL

References:

[1] Jingook Kim, Jongjoo Lee, Seunyoung Ahn; "Closed-Form Expressions for the Noise Voltage Caused by a Burst Train of IC Switching Currents on a Power Distribution Network ", IEEE Transactions on Very Large Scale Integration (VLSI) Systems

[2] https://semiengineering.com/hbm-issues-in-ai-systems/

[3] Amit Paunikar, Saurabh Arya, Vikas Makhija, Shaily khare; " Don't delay catching bugs: Using UVM based architecture to model external board delays"

[4] JESD238A-High Bandwidth Memory JEDEC Specification