SAND2021-8650C

Towards information limit in a low photon 3 dimensional x-ray imaging

Jungki Song*, Zachary Levine, Zhen Guo*, Bradley Alpert***, Michael Glinsky****, and George Barbastathis***

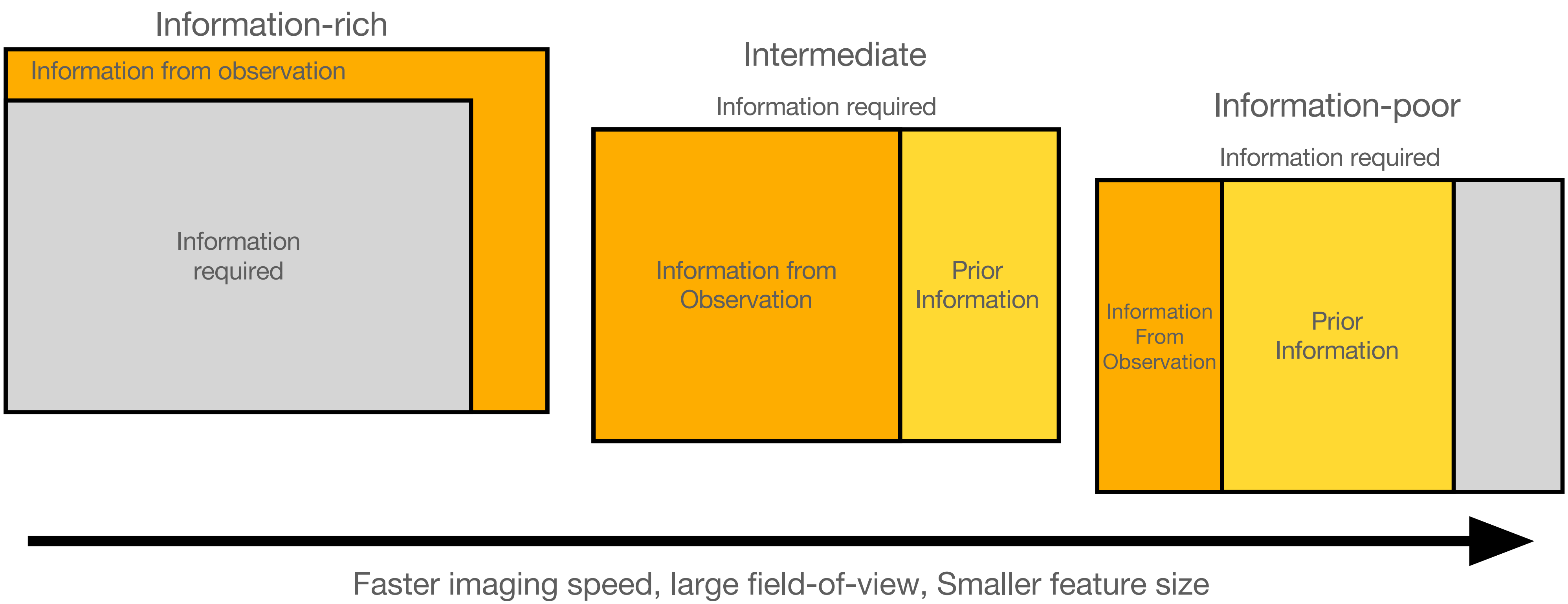
***MIT**
****NRL**
*****NIST**
******Sandia National Laboratories**

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Information-based approach towards imaging

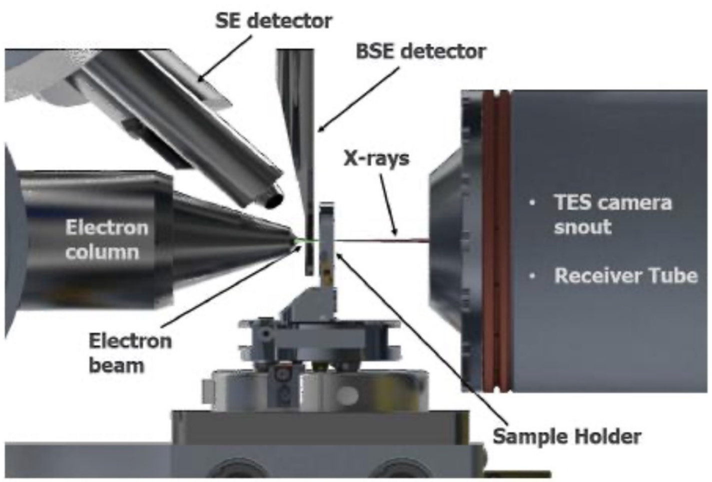
Classification of imaging problems based on information amount

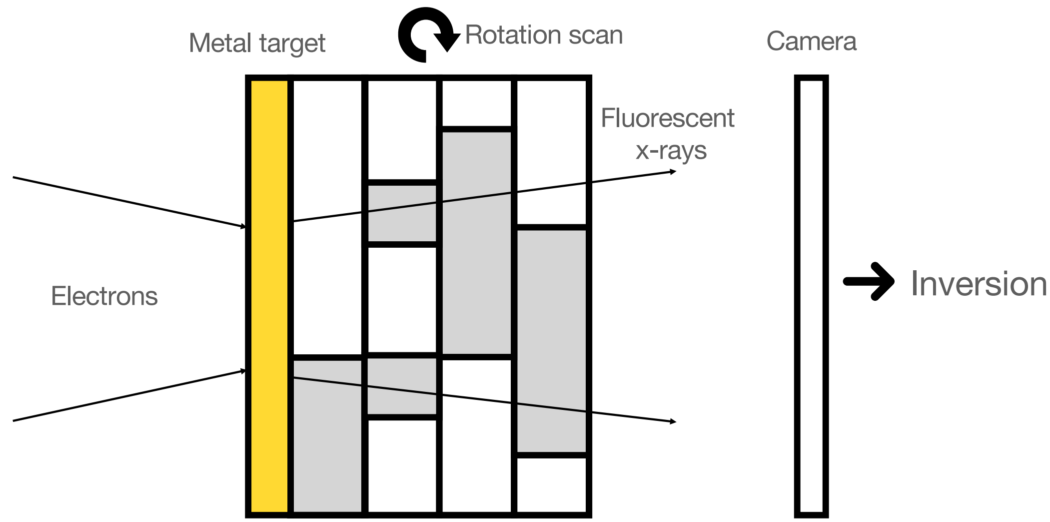


Collaboration between
MIT, NIST, and Sandia National Laboratory

Overview

Develop inverse algorithms for x-ray imaging 3D integrated circuits





Questions to address

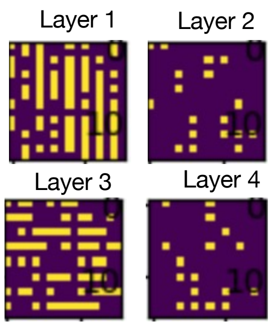
- How much information is required for 3D imaging?
- How do machine learning algorithms perform compared to classical maximum-likelihood estimation algorithm?
  - information-rich, intermediate, and information-poor
- What is an information-theoretic bound for imaging, and how can we compute it?

Collaboration between
MIT, NIST, and Sandia National Laboratory

Research routine

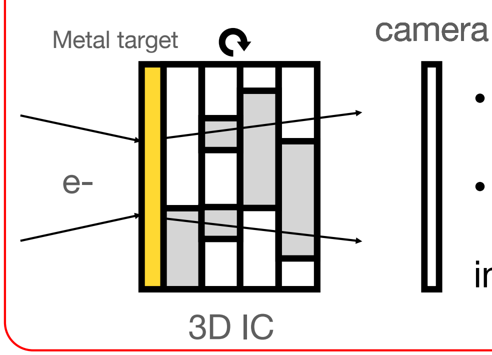
Generate fake 3D circuits

CircuitFaker



- Generate binary circuits
- Compute entropy of the circuit

Simulate x-ray imaging of 3D integrated circuits



- Tomographic measurements
- Control photon budget (Information-rich, intermediate, information-poor)

Acknowledgement: Dr. Courtenay Vaughn

Design inverse algorithms

- Maximum-likelihood (classical algorithm)
- Physics-assisted machine learning (PAML)
- Generative adversarial network (GAN)

Performance test (classical vs ML)

- Define evaluation criteria: bit error rate
- Compare classical vs ML algorithms
- Find information-rich, intermediate, and information-poor regime

Compute information bound

- Mutual information
- Monte Carlo approximation
- Theoretic bound in reconstruction quality at a given photon budget

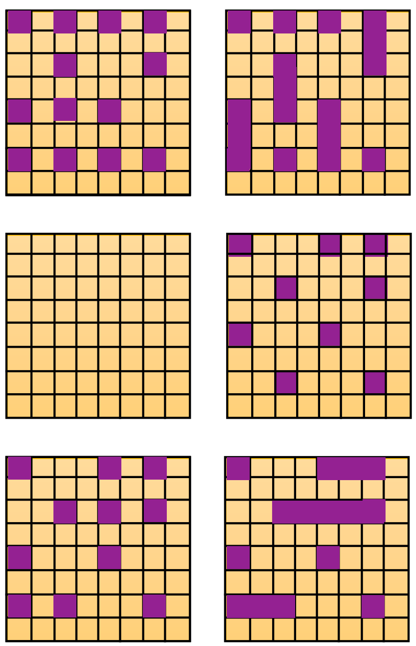
CircuitFaker: generate circuit with pre-defined design rule

- Generate binary circuits (Cu or SiO2) that emulate real-world integrated circuits
- Can compute amount of information within the circuit

Layer 1
(vertical wire)

Layer 2
(vias)

Layer 3
(horizontal
wire)

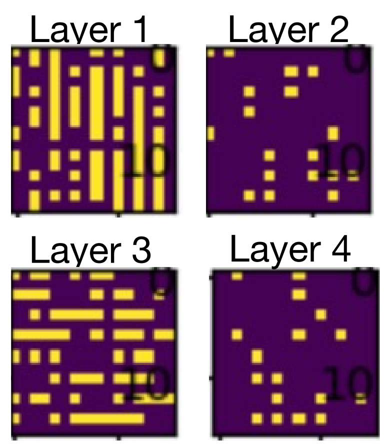


Sequence

Step 1: Seed layers

Step 2: Wire the odd layers
vertically/horizontally

Step 3: Connect vertically with vias



N Bernoulli trials

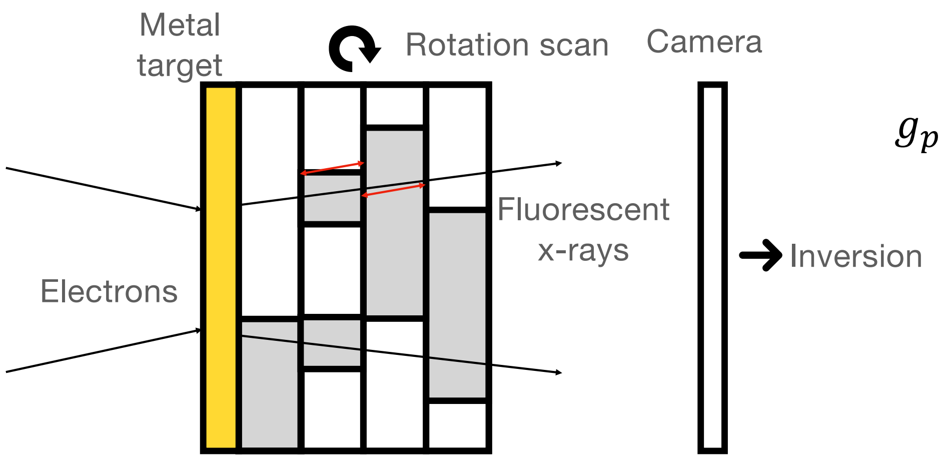
m_j Bernoulli trials
(j=x,y,z)

Example computation

For 16 x 16 x 4 circuits according to circuitFaker

Information required	185.5 bits

Simulate x-ray imaging of 3D integrated circuit



Attenuation of x-ray (no scattering)

$$g_p = Poisson(N_0 e^{-Af}) \Rightarrow log(N_0/g_p) = Hf$$

Linear system

H: ray propagation operator (distance the ray travels inside each voxel)
f: 3D circuit (attenuation coefficient)
N_0: Initial number of photons per ray
g_p: Poisson-contaminated measurements

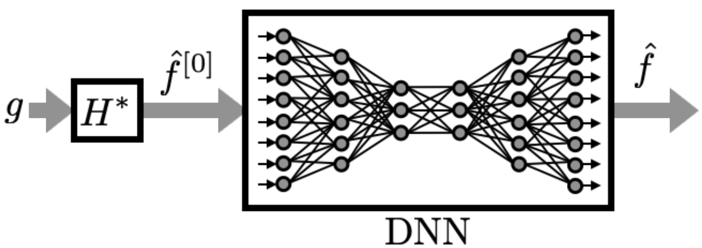
Imaging conditions

- Reduce angular scan range (reduce dimensions of g_p)
- Reduce photon budget (N_0)

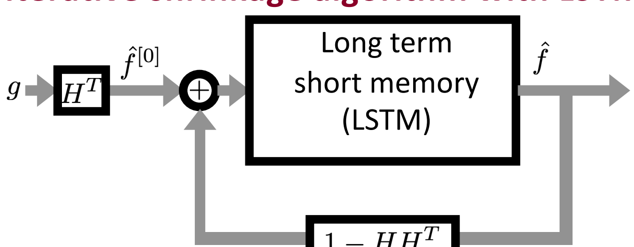
Machine learning algorithms development

Approach 1: Physics-assisted machine learning

a) Approximant-based



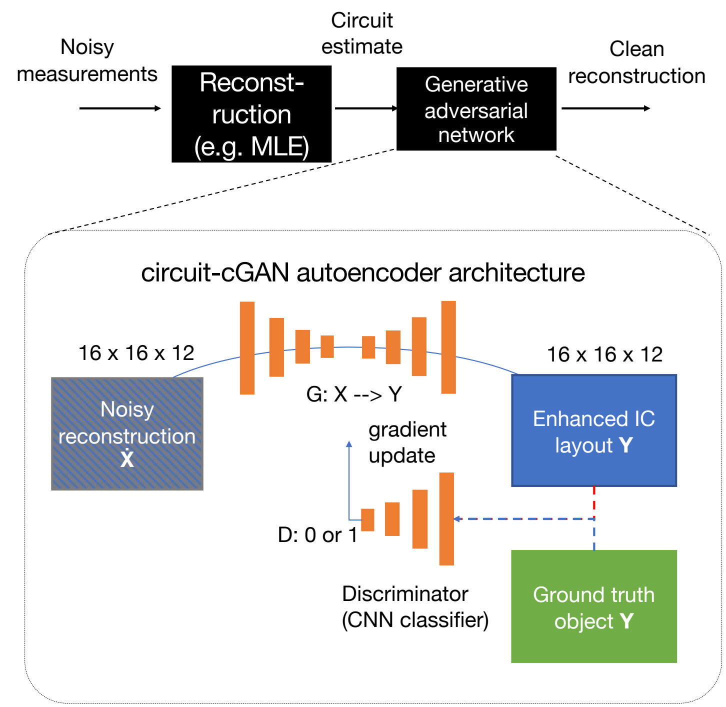
b) Iterative shrinkage algorithm with LSTM



$$\hat{f} = argmin_f \parallel \mathbf{Hf} - \mathbf{g} \parallel_2^2 + \alpha \parallel f \parallel_1$$

$$\hat{f}_{t+1} = Prox[(\alpha H^T log(N_0/g_p) + (I - \alpha H^T H)\hat{f}_t] = LSTM[(\alpha H^T log(N_0/g_p) + (I - \alpha H^T H)\hat{f}_t)]$$

Approach 2: Reconstruction + generative adversarial network (GAN) denoising



Bit error rate: evaluation criteria

Frequency of wrong predictions in classifying materials in IC voxels

Maximum likelihood classifier

Step 1: Compute $p(f=0|\hat{f})$ and $p(f=1|\hat{f})$.

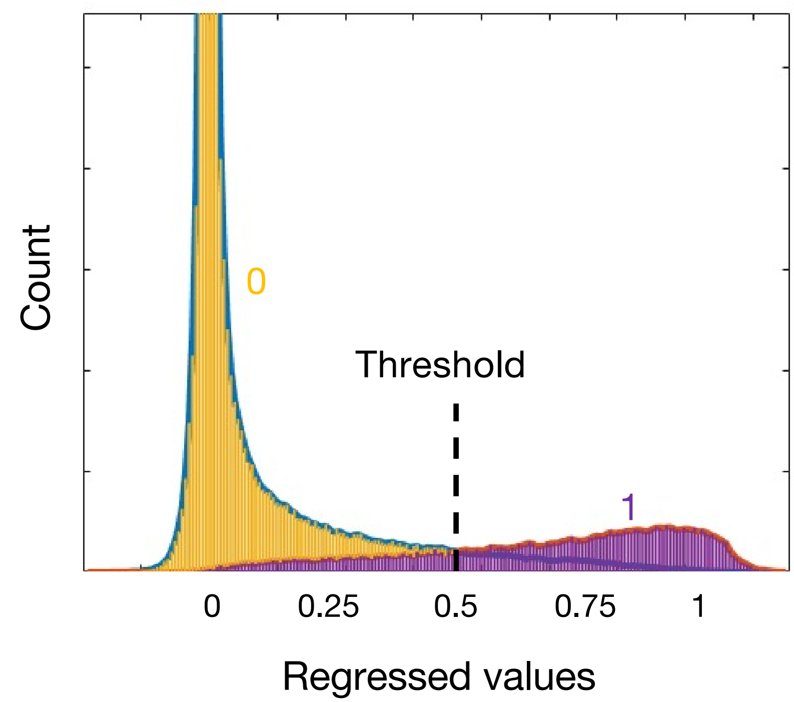
$p(f=0|\hat{f}) = p(\hat{f}|f=0)p_0$: Likelihood for 0

$p(f=1|\hat{f}) = p(\hat{f}|f=1)p_1$: Likelihood for 1

p_0, p_1: Prior distribution of 0 and 1 in ICs

$p(f=0|\hat{f}) > p(f=1|\hat{f})$: Classify as 0

$p(f=0|\hat{f}) < p(f=1|\hat{f})$: Classify as 1

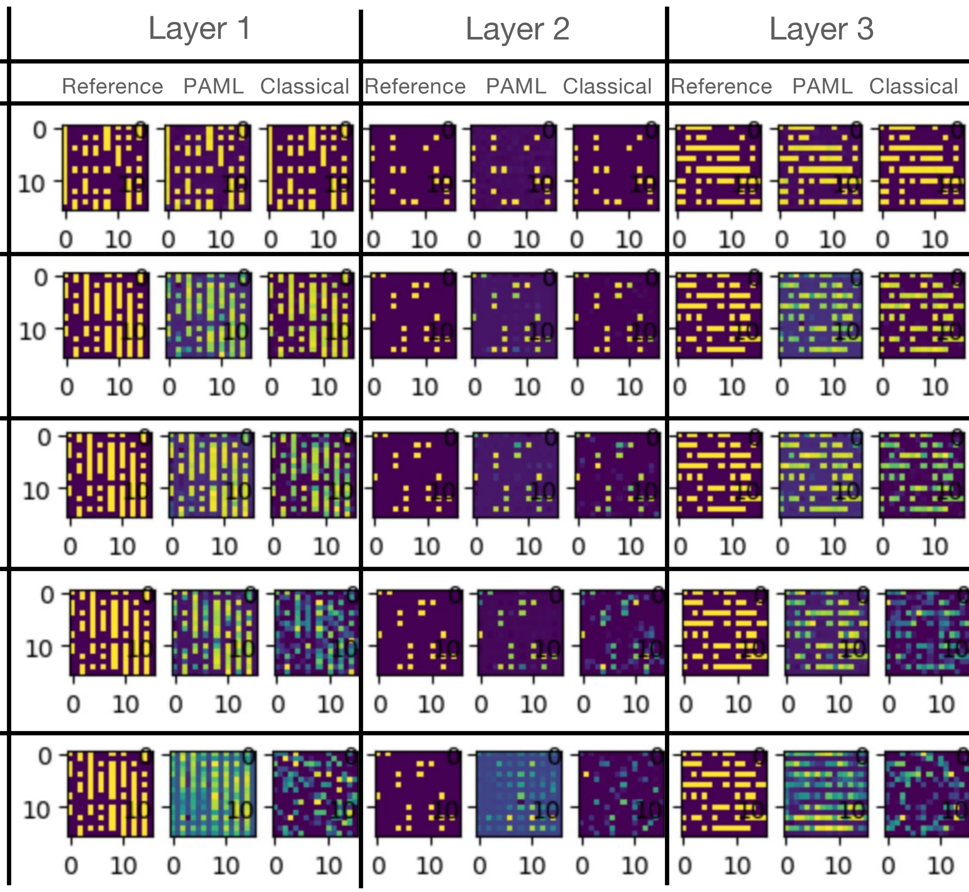


Step 2: Choose intersection point as a threshold in order to classify 0 & 1.

Step 3: Compute error rate for 0 & 1 (BER_0 and BER_1).

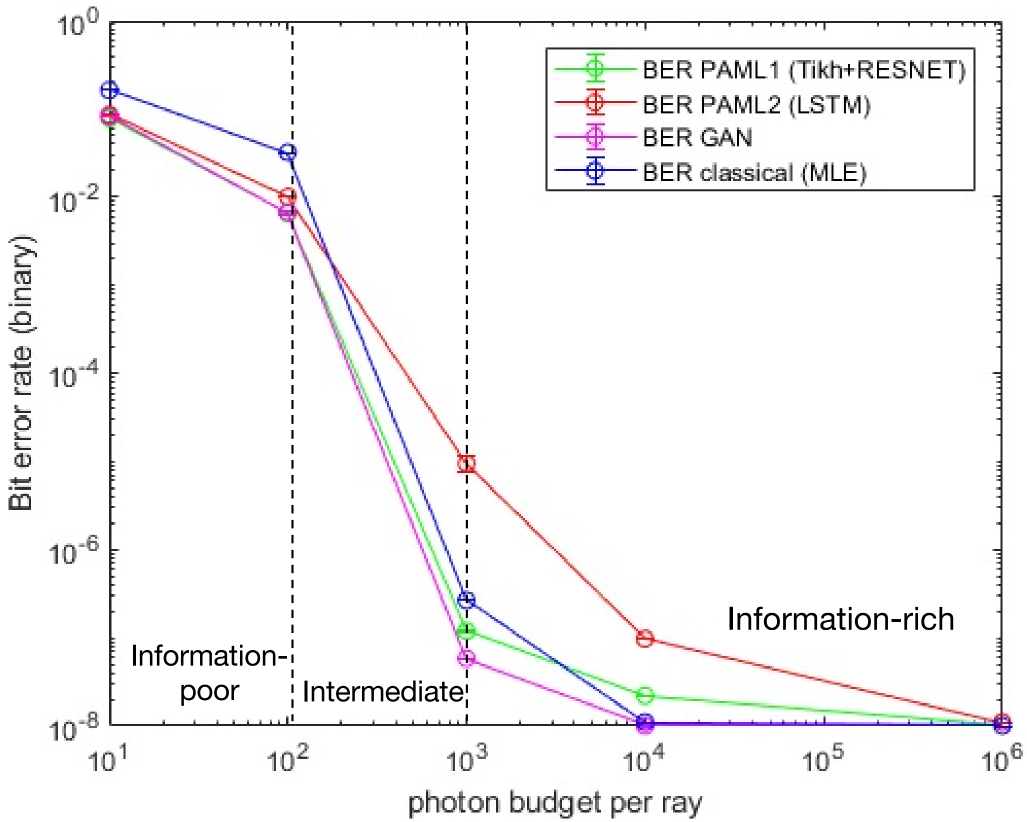
Reconstruction results

- 24 tomographic angles
- +/-90 deg



	Bit error rate (PAML / classical)
10^6 photons	1e-8 / 1e-8
10^4 photons	2.2e-8 / 1.1e-8
10^3 photons	1.2e-7 / 2.6e-7
10^2 photons	6.5e-4 / 3.2e-2
10^1 photons	7.7e-3 / 1.6e-1

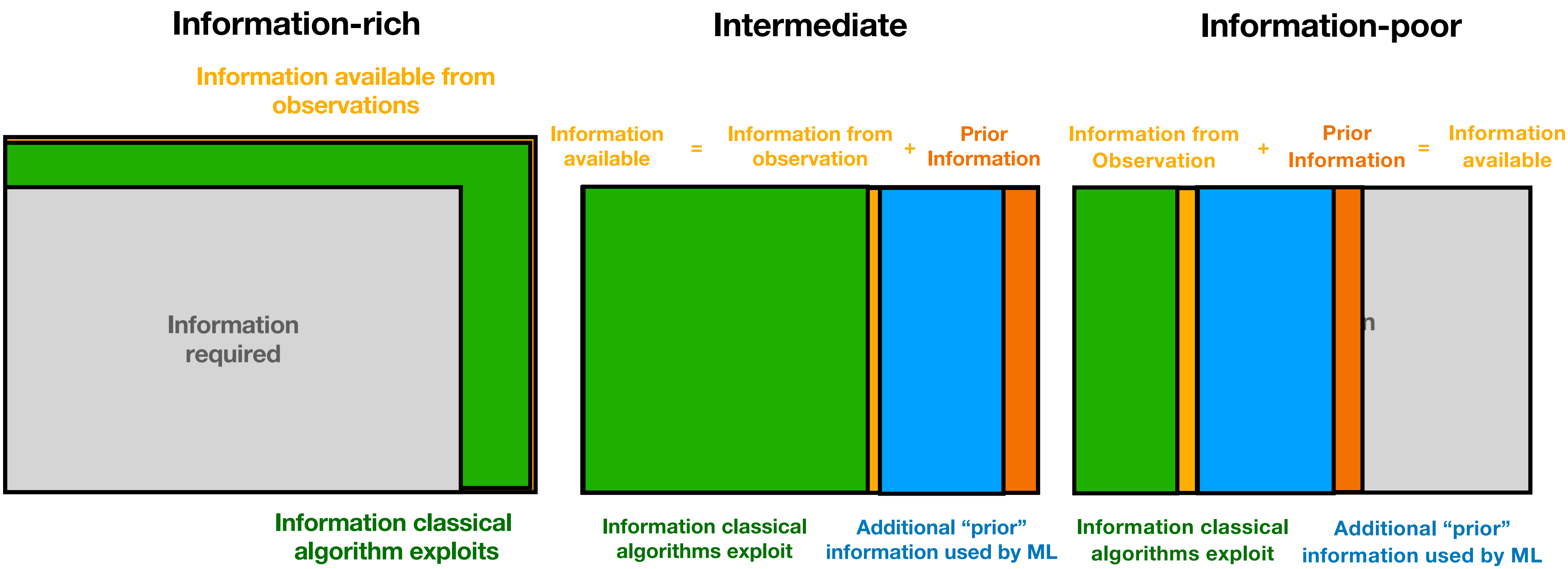
Performance summary



- **Information-rich**: classical maximum-likelihood and machine learning (ML) algorithms show comparable performance.
- **Intermediate**: ML starts to excel. Compromise between accuracy and dwell time
- **Information-poor**: ML algorithms show ~5x improved error rates at low photon budget (<100).

Information-theoretic bound for imaging

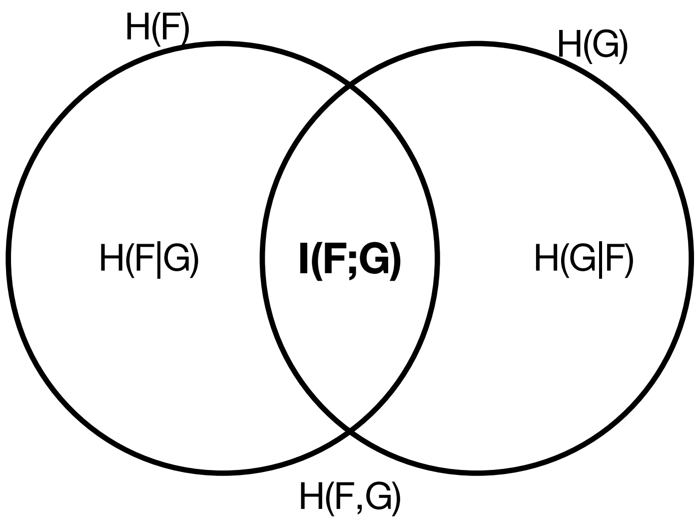
How much information can we retrieve from imaging?



How can we compute total amount of “available information”?

Information-theoretic bound for imaging

Mutual information



- $C = I(F;G) = H(F) - H(F|G)$
  $= I(G;F) = H(G) - H(G|F)$
- $H(F)$: Entropy, total amount of information to be retrieved in 3D circuits
- $H(F|G)$: Conditional entropy
  $H(F|G) = 0$: Perfect imaging (measurement G fully retrieves information in F)
  $H(F|G) > 0$: Imperfect imaging (measurement G can't fully retrieved information in F)

Residual uncertainty:

$$1 - \frac{I(F;G)}{H(F)}$$

$$= 1 - \frac{120.8}{185} = 0.35$$

Computation of $I(G;F) = H(G) - H(G|F)$

Monte Carlo approximation

$H(G|F)$

Known from CircuitFaker

H(G|F) = $\sum_f H(G|F=f)p(f)$

Approximate
With Monte Carlo

$= E_f[H(G|F=f)] \sim \frac{1}{N_{sample}} \sum_{f_{MCsample}} H(G|F=f)$

Conditional entropy for N independent Poisson pdf's
(N: pixel counts)

For loop:
Sample circuit, f
Compute conditional entropy $H(G|F=f)$
End
Average over $H(G|F=f)$ to compute $H(G|F)$

$H(G)$

Nested for-loops

Inner loop: compute p(g) with Monte Carlo

$p(g) = \sum_f p(g|f)p(f) = E_f[p(g|f)] \sim \frac{1}{N_{MC}} \sum_i p(g|f_i)$

Outer loop: compute H(g) with Monte Carlo

$H(G) = -\sum_g p(g) log p(g) = -E_g[log p(g)]$

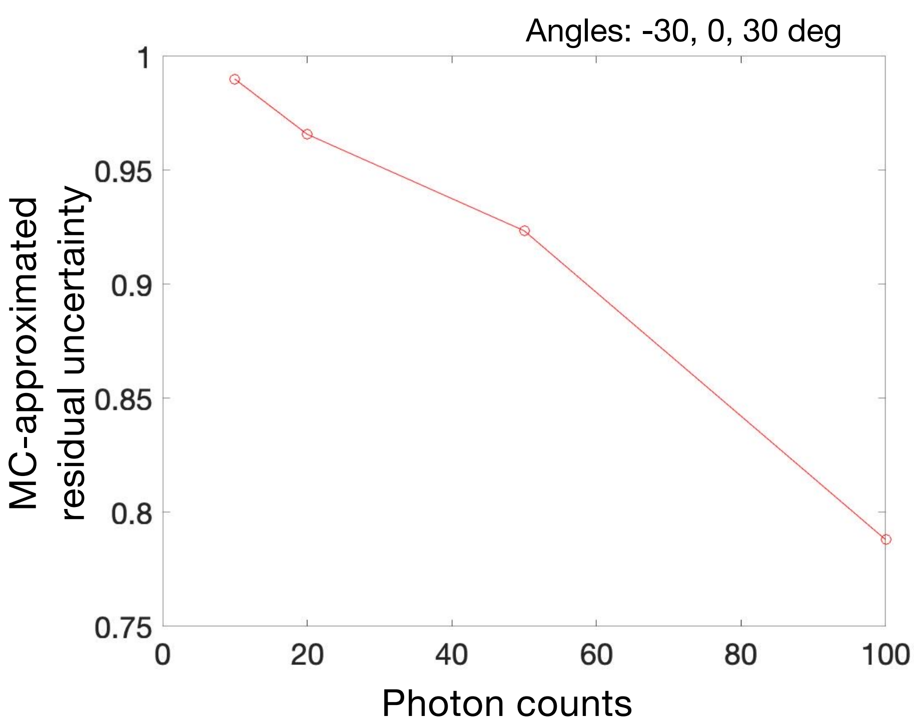
Outer for-loop:
Inner for-loop:
1. sample f
2. compute p(g|f)'s using sampled f

end
Average over p(g|f) to compute p(g) (Bayes' rule + MC integration)
end
Average of -log(p(g))'s to compute H(g) (MC integration)

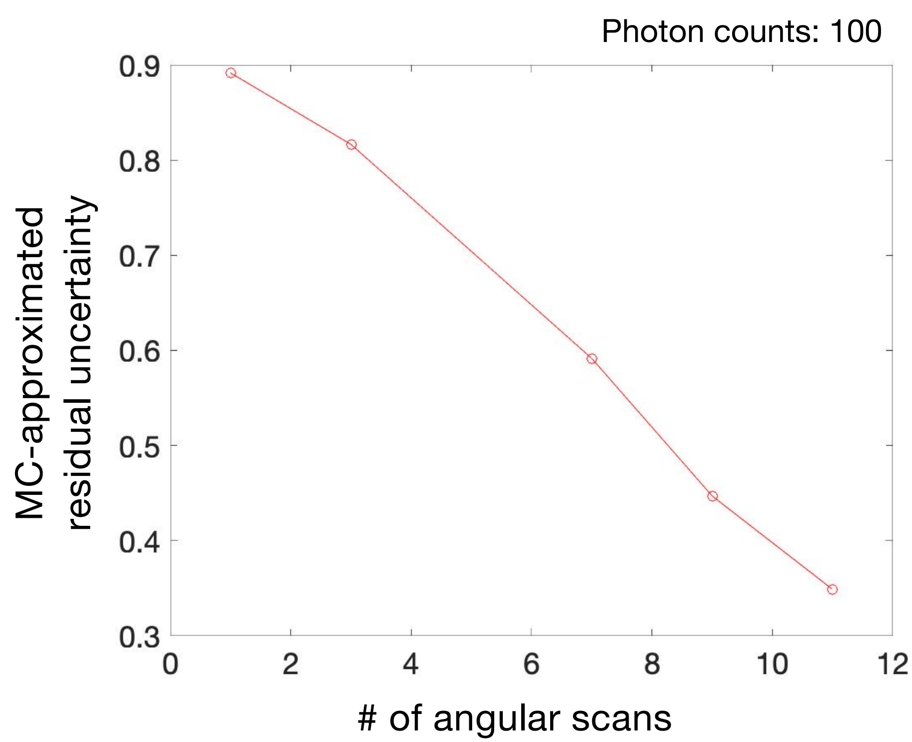
Example computation

Theoretical bound for BER in reconstruction of 16 x 16 x 4 circuits

1) Changing photon counts



2) Changing number of scan



Summary

- Categorized imaging problem according to amount of information available: information-rich, intermediate, and information-poor.
- In information-rich regime, no gain from using ML compared to maximum-likelihood algorithm.
- In information-poor regime, ML excels maximum-likelihood algorithm by exploiting prior information, but not accurate enough for practical use.
- In intermediate regime, ML help reduce efforts in observations (e.g. reduce scan time) without compromising accuracy and practicality.
- Information-theoretic bound can be computed using mutual information and Monte Carlo approximation.