# Treaty Verification without an Information Barrier

Chris MacGahan, Ph.D. Candidate at University of Arizona

Advisor: Dr. Matthew Kupinski

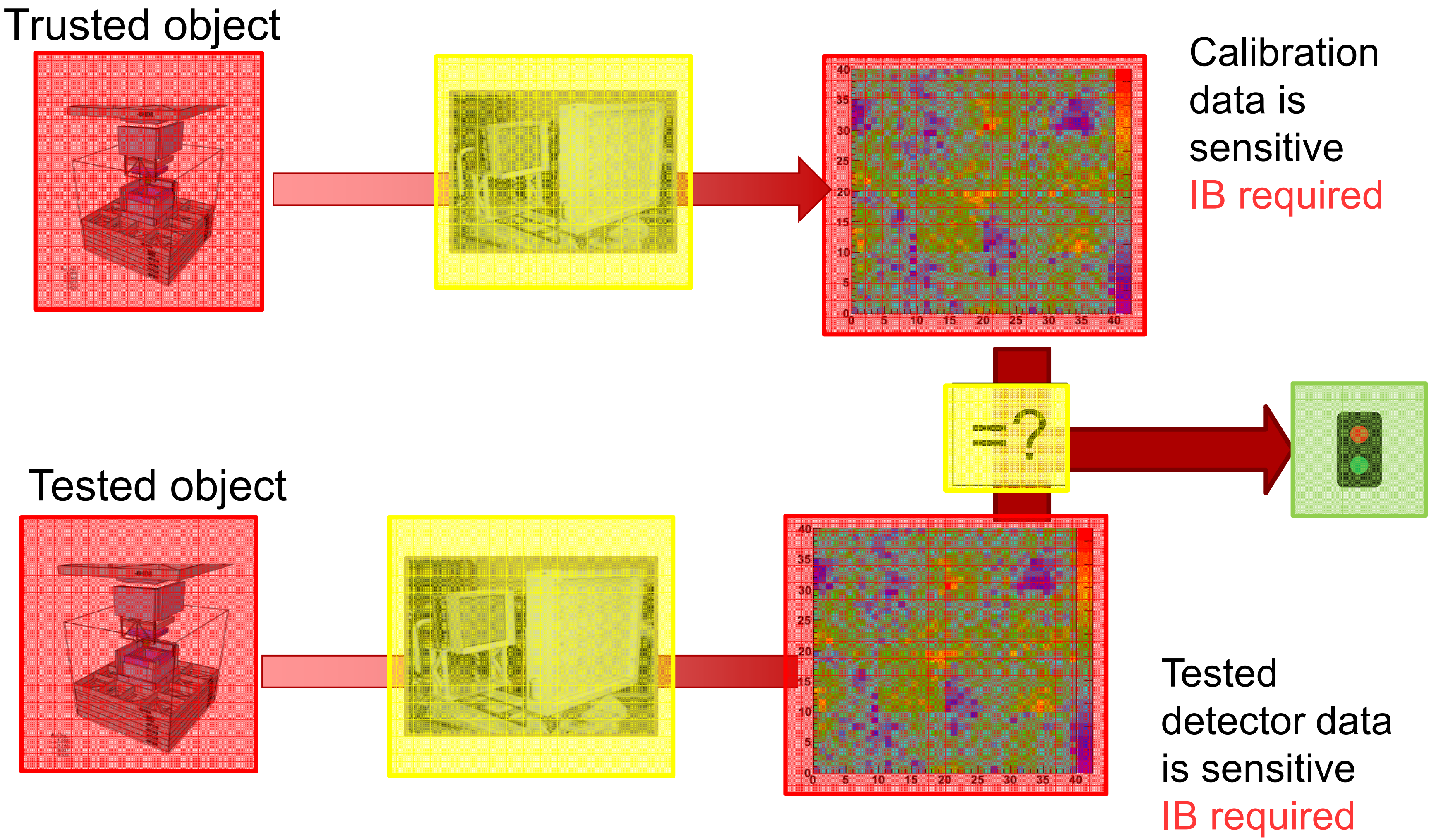
Mentors: Erik Brubaker & Nathan Hilton, Org: 8127

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

# The problem

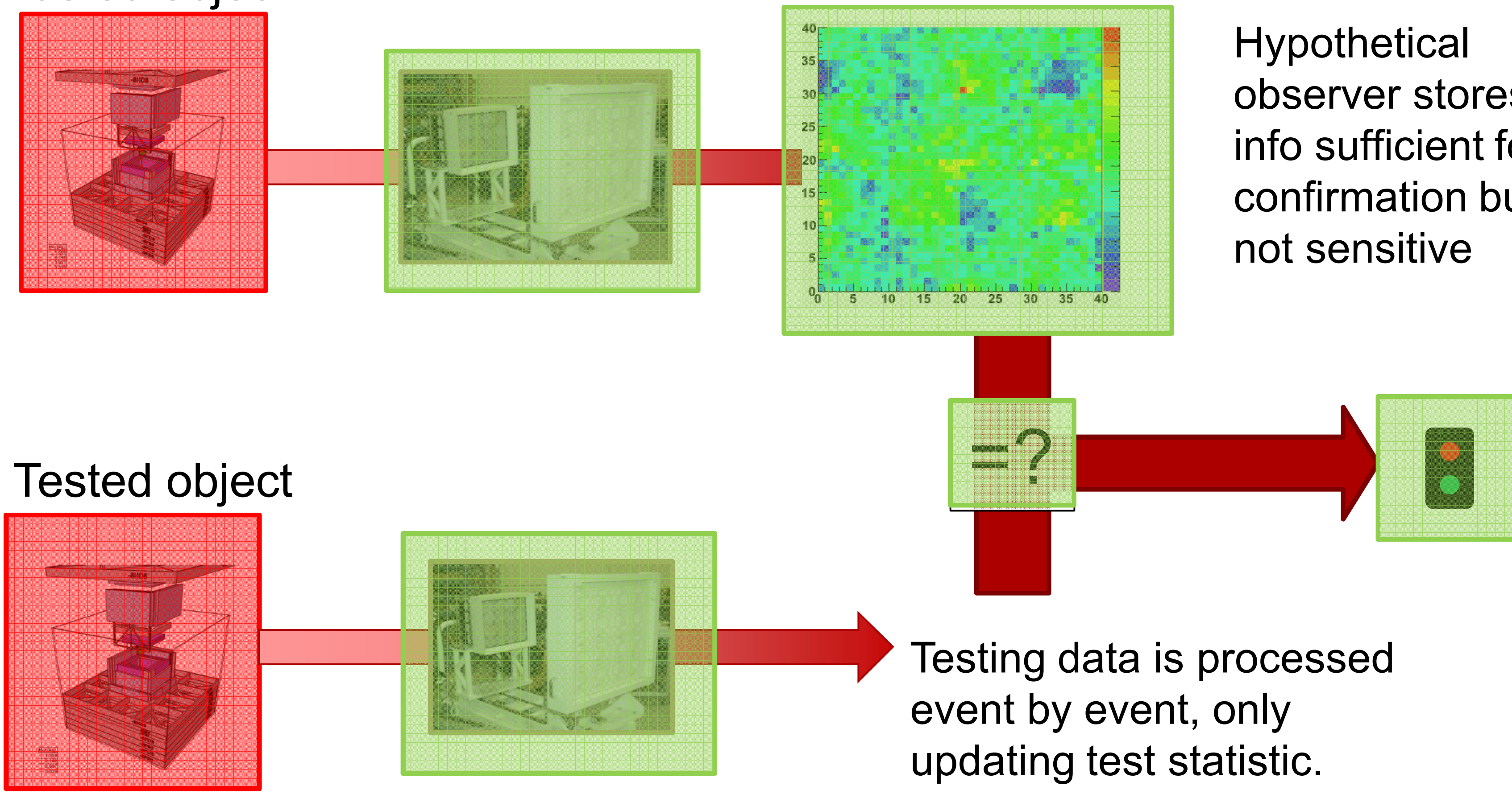
- Current treaty verification tests for delivery systems
- What if countries want to test if a warhead has been disarmed?
- Monitor wants to verify, host wants to preserve sensitive information on construction of objects.
- Many current proposed methods utilize an information barrier (IB)
  - IB: hardware or software

# “Traditional” template matching



# Our proposal

Trusted object




Hypothetical observer stores info sufficient for confirmation but not sensitive

Tested object

Testing data is processed event by event, only updating test statistic.

Data not aggregated

# Definitions

- Detector data can be described by number of total detected counts *N* and list-mode data $\{A_n\}$.

- $\{A_n\}$ contains the $n^{th}$ detected particle type, pixel #, energy bin.

- Randomness in the source - orientation, material age, construction, storage container. We call set of nuisance parameters that degrade performance $\gamma$

# Linear template observers

- Testing and training event data $\{A_n\}$ binned into data vector *g* (P x 1).

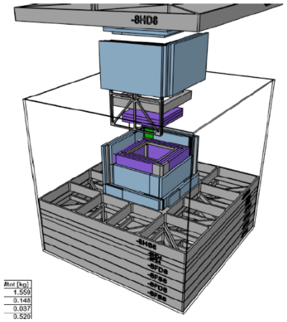
$$g_p = \sum_{n=1}^{N} f_p(A_n)$$

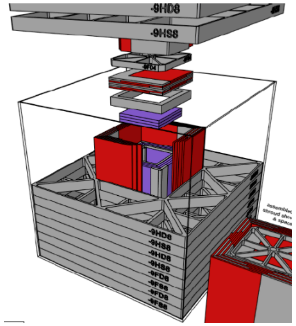
- Linear template W (P x 1) acts on $g_{test}$ , result is thresholded to make a decision

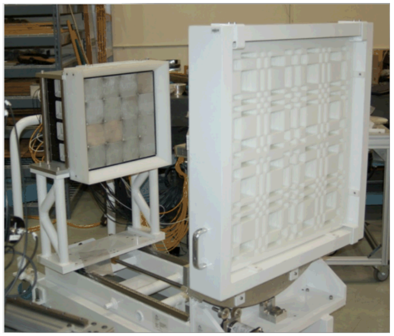
$$t_{test} = W^\dagger g_{test}, \; t_{test} \lesseqgtr t_{thresh}$$

# Experiment (simulation)

- Binary discrimination using spectral information.
  - Distinguish objects 8 (Pu surrounded by DU) and 9 (Pu surrounded by HEU) developed by Idaho National Lab.
  - Fast-neutron coded-aperture detector with liquid scintillator.



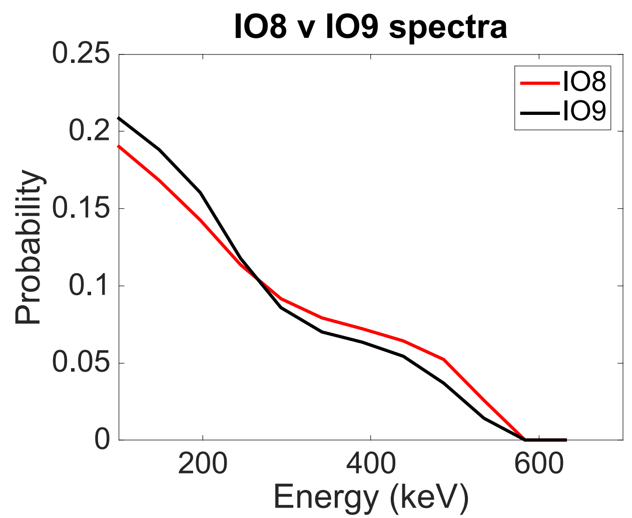




- Rotational variability included (simulated grid of orientations)
- Models built into transport application using GEANT4 toolkit to acquire testing and training data.

# Hotelling observer

- Data on the inspection objects differ in their spectra and count rate.

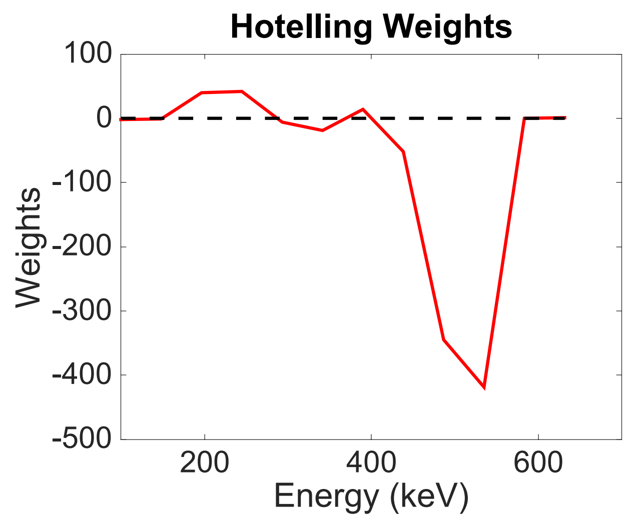



- Hotelling observer is template *W* defined as:

$$W = K_g^{-1}\overline{\Delta g}$$

$$K_g^{-1} = \frac{K_1 + K_2}{2}$$

$$\overline{\Delta g} = \overline{g_2} - \overline{g_1}$$



# Notes on nuisance parameters

- When incorporating source variability, data becomes doubly stochastic

$$\overline{\overline{g_1}} = \left\langle \langle g_1 \rangle_{g|\gamma} \right\rangle_\gamma$$

Averaged over Poisson variability (given known nuisance parameter), then source variability

$$K_1 = \left\langle \left\langle (g_1 - \overline{\overline{g_1}})(g_1 - \overline{\overline{g_1}})^\dagger \right\rangle_{g|\gamma} \right\rangle_\gamma$$

$$K_1 = \langle K_{1,n}(\gamma) \rangle_\gamma + K_\gamma$$

- On left is Poisson covariance (equal to number of observed counts) integrated over source randomness
- On right is covariance of data due to source variability
- Easy to invert via Matrix Inversion Lemma

# Storage for Hotelling observer

- How sensitive is the stored information that our observer model uses?

- Template W contains product of first and second order statistics, but still (likely) constitutes sensitive information

# Channelized Hotelling

- Channelize vector *g*(Px1) with operator *T*(QxP) into much smaller vector *v*(Qx1) with *Q* values.

$$v = Tg$$

$$W_v = K_v^{-1}\overline{\Delta v}$$

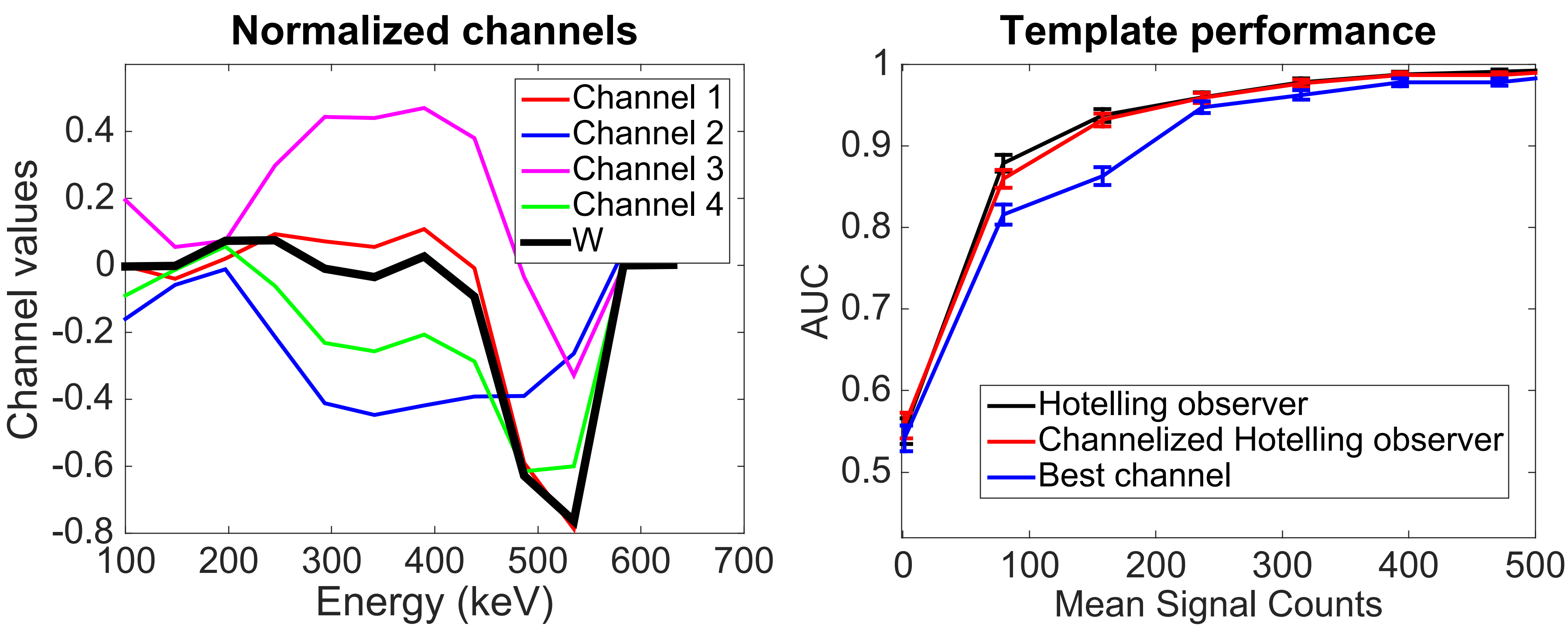
$$W_v^{\dagger} v_{test} \lessgtr t_{thresh}$$

- *T* can be optimized to maximize $SNR^2$ of test statistic distributions for best performance.
  - Simple gradient descent algorithm

$$f_{obj} = \overline{\Delta v}^{\dagger} K_v^{-1} \overline{\Delta v} = (T\overline{\Delta g})^{\dagger}(T^{\dagger} K_g^{-1} T)(T\overline{\Delta g})$$

# Channelized Hotelling regularizer

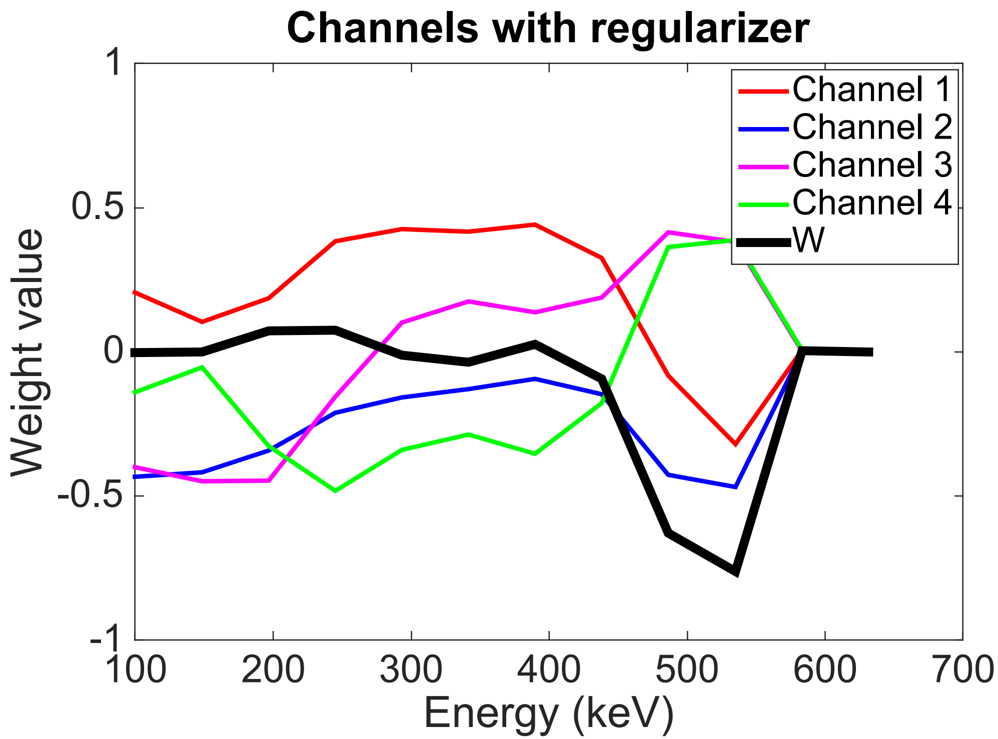
- However, standard optimization led to single or multiple strong performing channels. No point if channels are sensitive.



# Channelized Hotelling regularizer

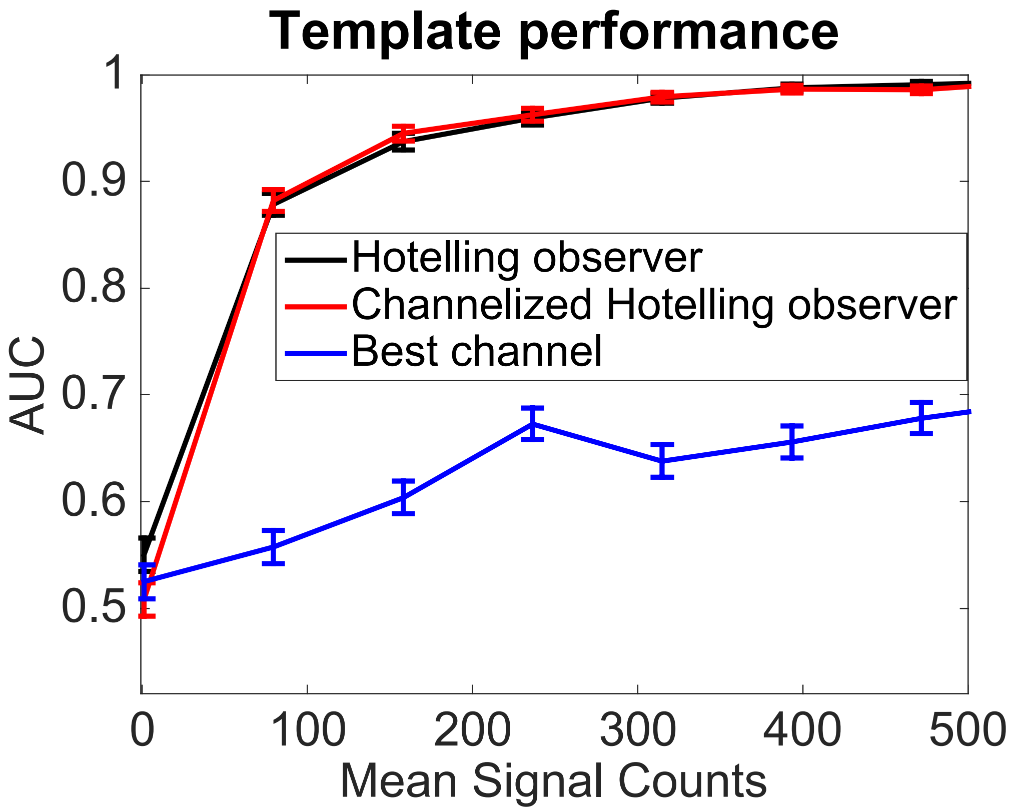
- Use regularizer in optimization to limit channel performance

$$f_{obj} = SNR^2_{Qchannels} - \eta \sum_{q=1}^{Q} SNR^2_{channel\ q}$$



# Channelized Hotelling regularizer

- Hotelling and channelized Hotelling perform well
- Individual channels perform poorly



# Summary

- Taken sensitive Hotelling template $W$ to $T$ and $W_v$, neither are sensitive without other.

- There is one caveat:

$$W = W_v T$$

If monitor is able to obtain access to both T and $W_v$, they can find sensitive Hotelling weights.

# Future work – null hypothesis test

- Binary classification is inherently spoofable.
- Need an observer to answer “Is this source A or not source A?”
- We developed a model based on likelihood expression, but it is spoofable.
- Standard tests based on distance metrics
- Is there a linear model similar to the Hotelling observer?

# Future work – reducing sensitive info

- Example: Source A is a BeRP ball with 1” of poly shielding. The host country doesn’t want the monitor to know what source A’s poly thickness is down to a tolerance of Δt

$$f_{obj} = SNR^2_{(B-A)} - \eta SNR^2_{(A_{(1" + \Delta t)} - A_{(1")})}$$

- Will lead to drop in performance with benefit that host needn’t worry.

# Future work – reducing sensitive Info

$$f_{obj} = SNR^2_{(B-A)} - \eta SNR^2_{(A_{(1"+\Delta t)} - A_{(1")})}$$

- A channelizing matrix that optimizes this objective function wouldn't be based on sensitive data
- Likewise, sensitive data could not be gained through the inverse problem
- List-mode requirement would no longer exist