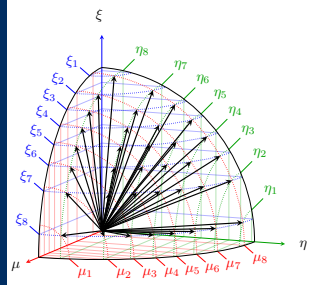
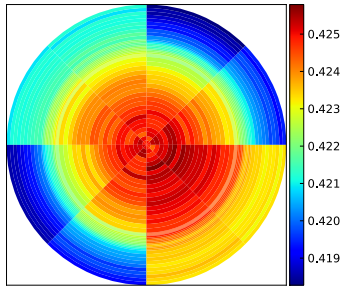
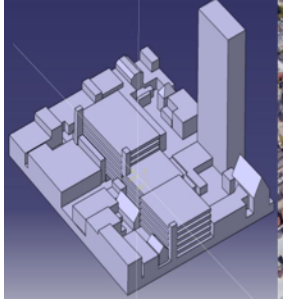






Uniformly Ordered Binary Decision Algorithm for Benchmark Experiment Correlations in Whisper Validation

Brian Kiedrowski
Nuclear Engineering and Radiological Sciences Department
University of Michigan
Los Alamos National Laboratory Subcontract

Nuclear Criticality Safety Program Technical Program Review
February 22-25, 2021

UNCLASSIFIED

Overview & Motivation

- Whisper is a statistical analysis package developed and maintained by LANL to support NCS validation

- The original methodology in Whisper does not account for benchmark experiment correlations when estimating the bias plus bias uncertainty (calculational margin)
  - Correlations between benchmarks reduce the amount of information in the population and its statistical predictive power

- The Uniformly Ordered Binary Decision Algorithm has been developed to address this shortcoming

Outline

- Summary of Original Whisper Algorithm
- Uniformly Ordered Binary Decision Algorithm
- Results
  - HEU Solutions
  - LEU Lattices

Whisper Overview

- Whisper uses MCNP6.2 sensitivity profiles with ENDF/B-VII.1 and and covariance data to compute c_k to assess similarity between benchmark experiment and application models

- The c_k similarity values are used to assign statistical weighting factors w_i for each benchmark
  - Weighting factors can be thought of as a probability of including a particular benchmark experiment in a hypothetical validation exercise
  - Assigned proportional to the similarity coefficients to ensure an adequate effective sample size

Whisper Overview

- Whisper ensures there is an adequate population or sample size in the benchmark set
  - The effective sample size is the expected (mean) population of benchmarks in a hypothetical validation exercise
  - In the uncorrelated case, this is the sum of the weighting factors
  - Requires the effective sample size is greater than some threshold
    - nominally 25, but increases if the maximum c_k is too low

Whisper Overview

- Whisper uses the weighting factors to compute an extreme-value distribution
  - Probability distribution for the maximum negative bias
  - Cumulative distribution is the product of the weighted cumulative distribution functions of the normally distributed biases for each benchmark
  - Calculational Margin is the 99% upper confidence limit of the extreme-value distribution

Handling Correlations

- This approach does not account for benchmark experiment correlations
- Questions about the method with correlations
  - How to assign weighting factors?
  - How to quantify an effective sample size?
- Done with the Uniformly Ordered Binary Decision Algorithm

Uniformly Ordered Binary Decision Algorithm

- Find correlated clusters in the population, and, for each cluster, determine if an hypothetical analyst would consider pairs of benchmark experiments as redundant
  - Randomly order the cluster with uniform probability
  - Start with the first benchmark and include with a probability equal to the weighting factor w_i
  - For the next benchmarks, include with a probability equal to their weight, but decide whether to treat them as redundant with more similar benchmarks previously in the list
  - If both are included and redundant:
    1. Assign the worst case bias (done automatically by including both with extreme-value distribution)
    2. Count these two benchmarks as a single benchmark toward the effective sample size
  - If not redundant, include as an independent benchmark
  - Repeat for all possible permutations

Uniformly Ordered Binary Decision Algorithm

- The adjust weighting factor is the probability that a benchmark experiment will be included as independent or non-redundant
  - The unadjusted weighting factor is the probability that it is included either way
- The effective sample size is now the expected (mean) number of *non-redundant* benchmarks in the population
  - This is the sum of the adjusted weighting factors
- Since the adjusted weight is less than the original, the effective sample size is smaller
  - Causes Whisper to expand its benchmark search

Example

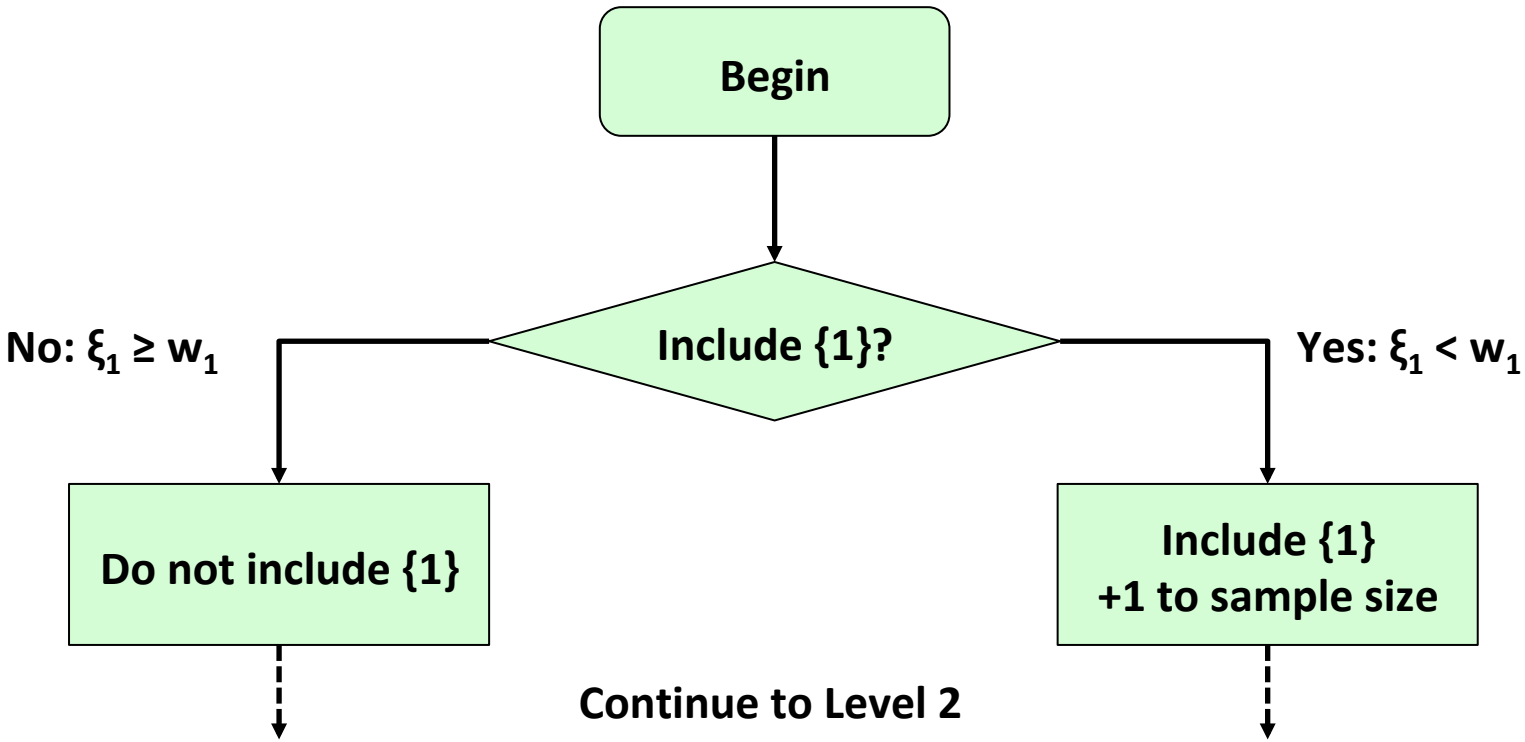
- Consider a cluster of 3 correlated benchmarks {1,2,3} in a random order

- Here we have three levels to the decision algorithm with each level giving the adjusted weight
  - p_j is the adjusted sample weight for the *j*th benchmark
  - w_j is the original or unadjusted sample weight
  - ρ_{ij} = is the redundancy probability:

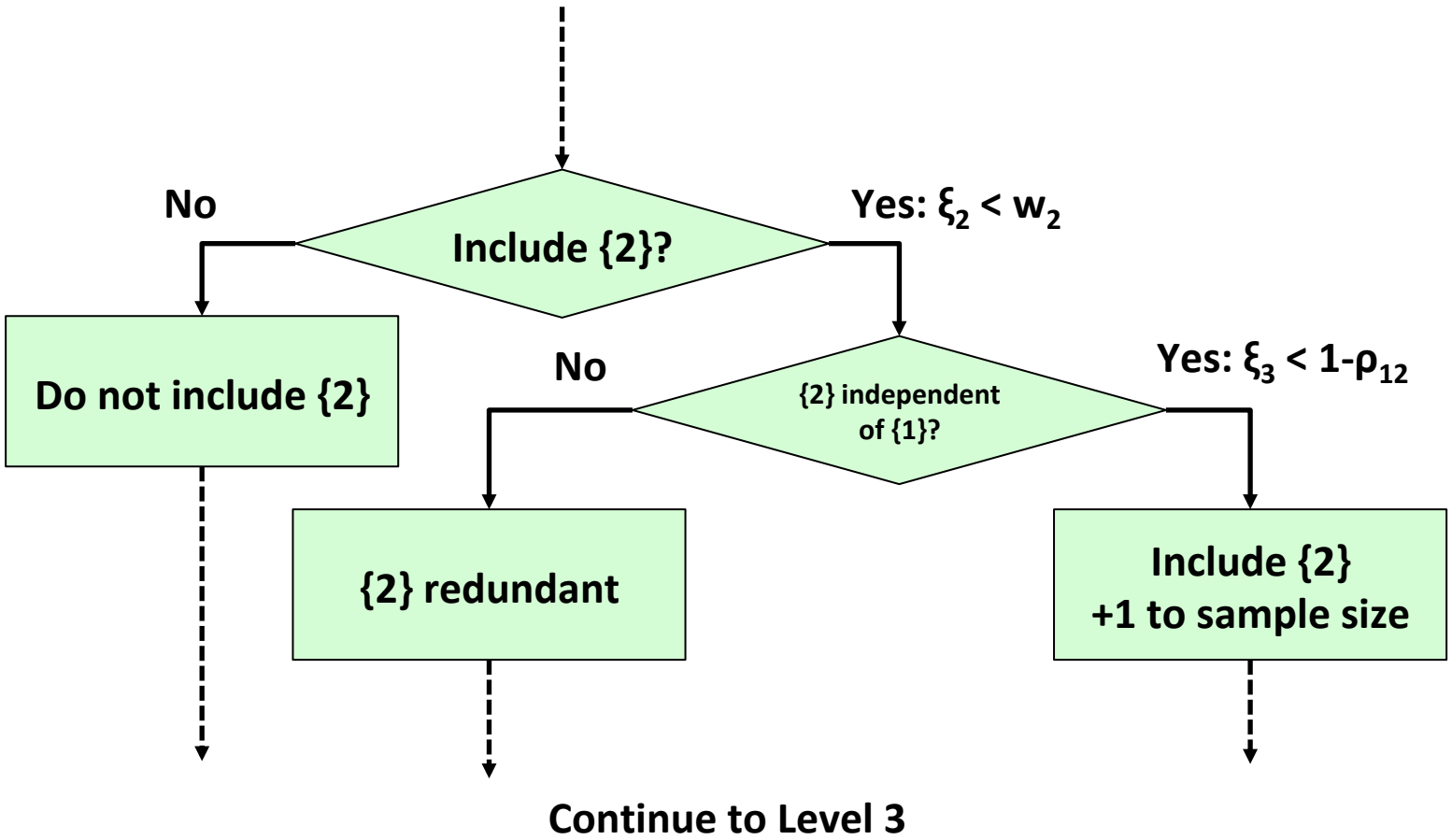
$$\rho_{ij} = r_{ij} \min\left\{1, \frac{w_i}{w_j}\right\}, \quad r_{ij} \geq 0$$

  - r_{ij} = benchmark correlation coefficient

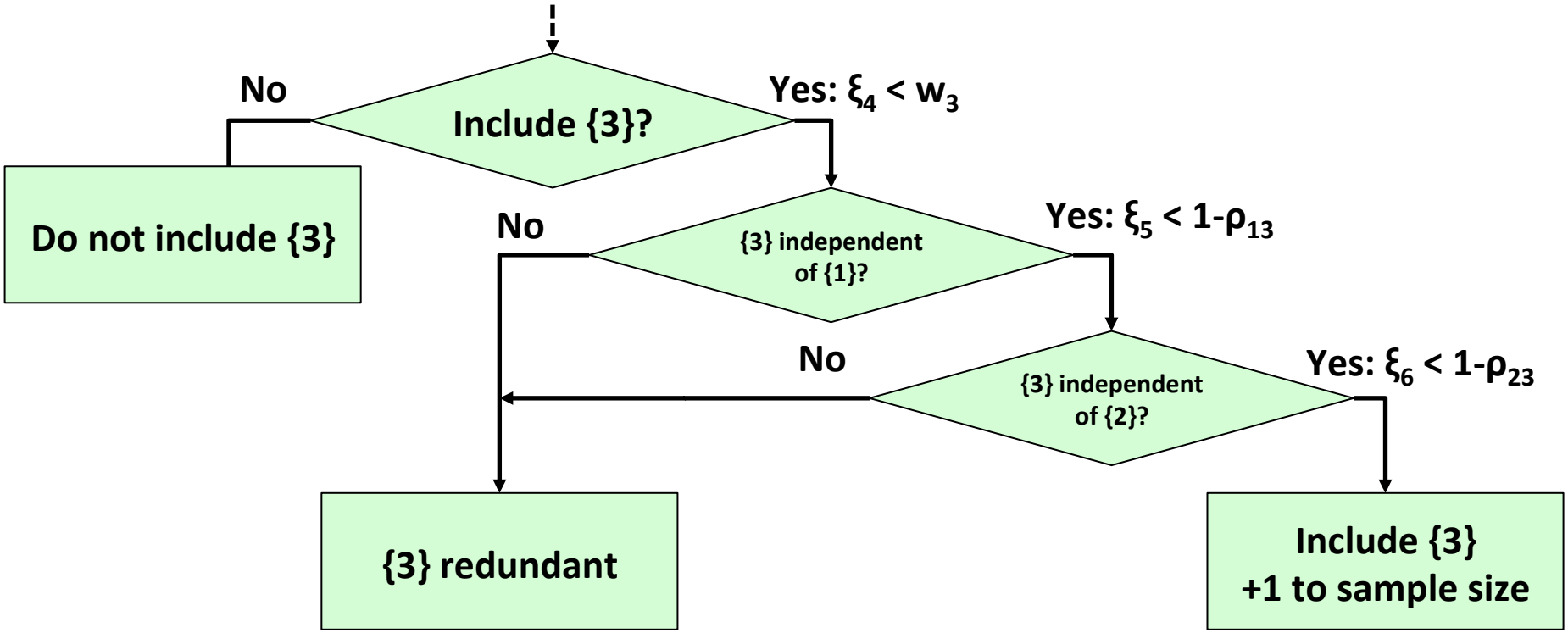
Example: Level 1



Example: Level 2



Example: Level 3



Adjusted Sample Weight

- We can deduce the probability for all permutations that a benchmark experiment is included as independent:

$$p_j = \frac{1}{n!} \sum_{\sigma} w_j \prod_{i=1}^{\sigma_j - 1} (1 - \rho_{ij})$$

  - n = cluster size
  - σ = all possible permutations of the cluster
  - σ_j = the position of *j* within the permutation
- The effective sample size is the sum of the p_j

Adjusted Sample Weight

- The summation in the expression for the adjusted sample weight contains $n!$ terms
  - Computationally tractable to compute directly for cluster sizes $n \leq 10$ on a modern computer
  - For larger clusters, use Monte Carlo sampling with random permutations and estimate

Effective Sample Size

- The computed effective sample size with correlations is compared against the sample size requirement
  - If effective sample size falls below requirement, decrease the acceptance criterion on c_k expanding the sample population, compute new weights, and repeat until the requirement is satisfied

- Once sample size is met, use the ***unadjusted*** weighting factors to find the extreme value distribution and calculational margin
  - Ensures the calculational margin is assigned conservatively

Results

- Implemented Uniformly Ordered Binary Decision Algorithm in a research version of Whisper1.1
  - Deliverable for the subcontract

- Transport calculations performed with MCNP6.2 and ENDF/B-VII.1 nuclear data

- Whisper calculations use data from Whisper1.1 library (except where otherwise noted)

Results

- HEU Solutions
  - Use application models identical to Cases 1-10 for HEU-SOL-THERM-001
  - Correlations available in DICE
  - Cases 9 and 10 are excluded by χ^2 rejection algorithm

- LEU Lattices
  - Use application models identical to Cases 1-3 of LEU-COMP-THERM-007 and Cases 1-10 of LEU-COMP-THERM-039
  - Correlations obtained from Scenario A from W.J. Marshall's doctoral dissertation (used to show impact of a highly correlated cluster)
  - Case 3 of LCT-007 is excluded by χ^2 rejection algorithm

HEU-SOL-THREM-001 Correlation Matrix

	HST-001-001	HST-001-002	HST-001-003	HST-001-004	HST-001-005	HST-001-006	HST-001-007	HST-001-008
HST-001-001	1.00	0.47	0.46	0.44	0.42	0.42	0.46	0.57
HST-001-002	0.47	1.00	0.42	0.58	0.42	0.42	0.41	0.44
HST-001-003	0.46	0.42	1.00	0.46	0.43	0.43	0.46	0.46
HST-001-004	0.44	0.58	0.46	1.00	0.44	0.42	0.42	0.44
HST-001-005	0.42	0.42	0.43	0.44	1.00	0.54	0.48	0.47
HST-001-006	0.42	0.42	0.43	0.42	0.54	1.00	0.48	0.47
HST-001-007	0.46	0.41	0.46	0.42	0.48	0.48	1.00	0.51
HST-001-008	0.57	0.44	0.46	0.44	0.47	0.47	0.51	1.00

HEU Solution Results

Application	# (orig.)	# (corr.)	$\sum w_i$	Δm
HST-001-001	55	68	34.5	0.0022
HST-001-002	52	62	31.0	0.0002
HST-001-003	49	69	34.7	0.0023
HST-001-004	53	66	31.2	0.0002
HST-001-005	41	72	34.1	0.0040
HST-001-006	43	73	34.0	0.0040
HST-001-007	49	69	34.6	0.0024
HST-001-008	51	69	34.6	0.0020
HST-001-009	53	66	31.4	0.0002
HST-001-010	41	72	34.2	0.0049

The presence of correlations causes the number of benchmarks included to increase.

The increase in the total sampling weight increases from 25 to around 30-35.

The calculational margin shows increases up to a few 100 pcm. This is because more benchmarks must be included.

HST-001-010 Application Analysis

Cluster	$\sum w_i$	$\sum p_i$	Benchmark	w_i	p_i
1	5.13	2.94	HST-001-006	1.000	0.333
			HST-001-005	0.996	0.330
			HST-001-003	0.877	0.273
			HST-001-007	0.875	0.261
			HST-001-008	0.868	0.248
			HST-001-001	0.861	0.257
			HST-001-002	0.520	0.143
			HST-001-004	0.506	0.138
2	4.97	3.35	HST-010-001	0.972	0.755
			HST-011-001	0.947	0.485
			HST-001-002	0.947	0.485
			HST-043-001	0.870	0.467
			HST-009-003	0.742	0.440
			HST-009-002	0.453	0.216
			HST-009-001	0.126	0.047
			HST-043-002	0.082	0.043
3	3.85	1.38	HST-025-005	0.888	0.339
			HST-025-004	0.868	0.298
			HST-025-001	0.857	0.290
			HST-025-002	0.857	0.290
			HST-019-001	0.384	0.166

There are three clusters of correlated benchmarks in a sample population including 72 benchmarks

The second and third columns show the reduction in the effective sample size because of correlation for each cluster

As expected, the adjusted sample weight or inclusion probability decreases proportionate to the weight and size and degree of correlation within the cluster.

LEU-SOL-THERM-007 and -039 Case Correlation Matrix

	LCT-007-001	LCT-007-002	LCT-007-003	LCT-039-001	LCT-039-002	LCT-039-003	LCT-039-004	LCT-039-005	LCT-039-006	LCT-039-007	LCT-039-008	LCT-039-009	LCT-039-010
LCT-007-001	1.00	0.93	0.39	0.98	0.98	0.97	0.97	0.96	0.96	0.97	0.97	0.98	0.97
LCT-007-002	0.93	1.00	0.56	0.92	0.92	0.92	0.93	0.92	0.93	0.93	0.92	0.94	0.94
LCT-007-003	0.39	0.56	1.00	0.40	0.39	0.41	0.42	0.46	0.46	0.42	0.39	0.43	0.39
LCT-039-001	0.98	0.92	0.40	1.00	0.98	0.97	0.97	0.96	0.96	0.98	0.97	0.97	0.97
LCT-039-002	0.98	0.92	0.39	0.98	1.00	0.97	0.97	0.95	0.96	0.98	0.97	0.97	0.96
LCT-039-003	0.97	0.92	0.41	0.97	0.97	1.00	0.97	0.94	0.95	0.97	0.96	0.97	0.97
LCT-039-004	0.97	0.93	0.42	0.97	0.97	0.97	1.00	0.96	0.95	0.97	0.96	0.97	0.96
LCT-039-005	0.96	0.92	0.46	0.96	0.95	0.94	0.96	1.00	0.95	0.96	0.94	0.96	0.96
LCT-039-006	0.96	0.93	0.46	0.96	0.96	0.95	0.95	0.95	1.00	0.96	0.95	0.96	0.96
LCT-039-007	0.97	0.93	0.42	0.98	0.98	0.97	0.97	0.96	0.96	1.00	0.97	0.97	0.97
LCT-039-008	0.97	0.92	0.39	0.97	0.97	0.96	0.96	0.94	0.95	0.97	1.00	0.96	0.97
LCT-039-009	0.98	0.94	0.43	0.97	0.97	0.97	0.97	0.96	0.96	0.97	0.96	1.00	0.97
LCT-039-010	0.97	0.94	0.39	0.97	0.96	0.97	0.96	0.96	0.96	0.97	0.97	0.97	1.00

LEU Lattice Results

Application	# (orig.)	# (corr.)	$\sum p_i$	Δm
LCT-007-001	74	81	34.4	0.0007
LCT-007-002	47	50	30.3	0.0003
LCT-007-003	51	51	25.1	0.0000
LCT-039-001	72	76	34.9	0.0006
LCT-039-002	70	75	35.0	0.0007
LCT-039-003	58	70	35.2	0.0009
LCT-039-004	68	77	35.0	0.0008
LCT-039-005	53	55	34.1	0.0005
LCT-039-006	53	62	34.6	0.0005
LCT-039-007	72	81	34.7	0.0007
LCT-039-008	65	72	35.2	0.0009
LCT-039-009	65	73	35.2	0.0009
LCT-039-010	53	71	35.1	0.0008

Conclusions are much the same as for the HEU solutions.

The increase in calculational margin is more modest, which is because there were already high negative bias benchmarks in the initial population prior to applying correlations

LCT-001-001 Application Analysis

Cluster	$\sum w_i$	W_c	Benchmark	w_i	p_i
1	10.34	0.98	LCT-007-001	1.000	0.170
			LCT-039-001	0.990	0.086
			LCT-039-002	0.985	0.084
			LCT-039-007	0.981	0.084
			LCT-039-008	0.967	0.083
			LCT-039-003	0.947	0.082
			LCT-039-009	0.938	0.081
			LCT-039-004	0.937	0.081
			LCT-039-010	0.842	0.073
			LCT-039-006	0.772	0.067
			LCT-039-005	0.685	0.060
			LCT-007-002	0.303	0.027

The correlations in Marshall's Scenario A are very high, > 0.95 in many cases.

This causes the cluster of 12 benchmarks to become essentially worth 1 benchmark with most benchmarks being included as redundant most of the time.

The adjusted sample weights are significantly decreased because of the high degree of correlation.

Conclusions

- The uniformly ordered binary decision algorithm has been developed and implemented into a research version of Whisper1.1
  - This addresses the outstanding issue of not accounting for benchmark experiment correlations when estimating the calculational margin
- Results were collected for HEU solutions and LEU lattices
  - The results show the presence of benchmark correlations may significantly increase the calculational margin because Whisper must expand the benchmark population
- **This method will lead to lower predicted USLs for some applications**

Final Thoughts

- Need to get results for fast spectrum systems
  - Several of the ZPR/ZPPR series have correlation data associated with them, but Whisper1.1 does not have models
- Benchmark correlation data is currently limited and the results will change as more correlations are added
  - There is a need to generate this data for a wider variety of cases
  - Could apply some arbitrary correlation for cases that DICE identifies as potentially correlated but has no quantitative data
  - Would be helpful to have benchmark correlations for different cases quantified during the development of benchmark evaluations
- Allows for the inclusion of different benchmark models from multiple sites
  - Different modeling assumptions may lead to different answer
  - Could all be included and assumed perfectly correlated

Acknowledgments

The author would like to acknowledge the generous support provided by the Nuclear Criticality Safety Program and the technical staff at Los Alamos National Laboratory the facilitating this subcontract that made this research possible

Discussion

Email contact: bckiedro@umich.edu