Multicenter comparison of alpha particle measurements and methods typical of semiconductor processing

Jeffrey D. Wilkinson, Brett M. Clark, Richard Wong, Charles Slayman, Barry Carroll, Michael Gordon, Yi He, Olivier Lauzeral, Keith Lepla, Jennifer Marckmann, Brendan McNally, Philippe Roche, Mike Tucker, Tommy Wu

***Abstract*— Alpha counting measurement methods have been widely used in the semiconductor industry for many years to assess the suitability of materials for semiconductor production and packaging applications. Although a number of published articles describe aspects of this counting, a multicenter, comparative trial has not been carried out to assess the methodological accuracy of current methods. This paper reports on experience with a 9 center, international, round-robin style trial using a shared set of samples to quantify variability in alpha emission measurements. Four samples representing low and ultralow alpha materials were counted by each participating lab in a blinded trial. The consensus mean emissivity for low alpha material was estimated as 30.9 khr^{-1}-cm^{-2} with a range from 20.2 to 45.5, less than half of which can be attributed to counting uncertainty or other known sources of error. A strong correlation for replicate measurements within a lab was also observed supporting the conclusion that there are systematic variations in equipment or calibration among labs. Eleven of 23 measurements of ultralow alpha materials were within 1 standard deviation of the consensus mean and 7 were at or below background. The high level of counting uncertainty for these measurements is thought to be sufficient to mask any systematic variation similar to the low alpha observations. Comparison of the reported values with a standard calculation demonstrates that there are also differences in the interpretation of the values reported for emissivity and error, underscoring the need for careful interpretation of results.**

Keywords-component; soft errors, alpha particles, semiconductor device reliability, device packaging

I. INTRODUCTION

Radioactive impurities in integrated circuit (IC) packaging materials are an important contributor to a circuit's soft error rate (SER), particularly as the process geometry and operating voltage are scaled to smaller feature sizes. Multiple reports have stressed the importance of ultralow alpha materials and the potential for the SER to be dominated by alpha emissions even when these materials are utilized [1, 2].

In the semiconductor industry, alpha particle emission levels are most commonly measured using sensitive, large area gas proportional detectors. Through careful sample preparation and control of detector parameters, it is thought to be possible to measure alpha emission rates of materials down to 1 khr^{-1}-cm^{-2} (1 emitted particle per square centimeter per 1000 hours). For many alpha emitting isotopes this corresponds to concentrations in parts per trillion (PPT) or less [1, 2]. For short half-life isotopes, such as ^{210}Po the corresponding concentration is as low as a few parts in 10^{18}. At these levels it is generally impractical to identify the particular isotope or combination of isotopes responsible for the alpha emission by chemical analysis. Clark [3] reports on an attempt to match uranium and thorium concentrations measured using glow discharge mass spectroscopic (GDMS) analysis with calculated and measured values of alpha emission. In all cases, the calculated emission rate based on the GDMS analysis underestimates the measured alpha emission flux by a factor between $4.2x10^2$ and $1.5x10^5$. Although assumptions of the particular alpha emitters and their secular equilibrium status have been made in previous work (see, for example [2, 4]) it is the authors' opinion that this has not been adequately demonstrated for these materials and would be particularly challenging due to the extremely low decay rates. This position is consistent with the guidance provided in the soft error test standard JESD89A ([5] page 76). To further complicate matters, there is no accepted flux standard for a large area alpha particle source with an emission rate within 1000X of the expected activity level. Likewise, there is no published standard for low-level alpha measurements, sample handling or laboratory quality control procedures, although a JEDEC standard is in preparation [6]. To better understand and quantify alpha counting measurements a loose consortium has been formed to coordinate joint projects. This paper reports on the results of the consortium's first project which is an effort to quantify the range of measurements made by different labs. This study did not attempt to separate different potential sources of variation, preferring to let "the chips fall where they may."

J.D. Wilkinson, Medtronic, CRDM Device Technology, Mounds View, MN, USA, jeff.wilkinson@medtronic.com, 763-526-0483. Inquiries related to the alpha consortium may be directed to Jeff Wilkinson, Brett Clark, Richard Wong or any of the other authors.

B.M. Clark, Honeywell, Spokane, WA, Brett.Clark@Honeywell.com, 509-252-8716 – R. Wong, Cisco Systems – C. Slayman, Ops A La Carte – B. Carroll, Freescale Semiconductor Inc. – M. Gordon, IBM TJ Watson Research Center, Yorktown Heights, NY USA, gordonm@us.ibm.com. – Y. He, Intel Corporation, CH5-157, ATTD, Chandler, AZ 85226 USA. – O. Lauzeral, iRoC Technologies – K. Lepla, Teck Metals Ltd. – J. Marckmann, Medtronic Tempe Campus, Tempe, AZ – B. McNally, XIA LLC – P. Roche, STMicroelectronics – M. Tucker, Alpha Sciences, Inc. – T. Wu, SGS Taiwan Ltd, Material and Engineering Lab, Polymer Lab, Kauhsiung, Taiwan, tommy.wu@sgs.com, +886-7-3012121.

This work was sponsored by funding from Cisco Systems, Medtronic, and Honeywell, along with donations of counter time from all of the participant companies.

Nine participants representing Alpha Sciences; Freescale Semiconductor; Honeywell Electronic Materials; IBM TJ Watson Research Center; Intel; iRoC Technologies; Medtronic, Inc.; SGS Taiwan Ltd; Teck Metals and XIA LLC (iRoC Technologies and Alpha Sciences participated jointly) made measurements for this study. Each participant agreed to use their normal laboratory procedures to make at least 4 measurements of samples representing a low alpha (LA) and an ultralow alpha (ULA) material. Although one previous publication has presented counter efficiency measurements made at multiple sites [7], this is the first study to directly compare alpha counting measurements at multiple sites at emission levels pertinent to the semiconductor industry.

All of the counters used in this study detected ionization from alpha particles as they passed through a gas-filled counting chamber. For the proportional mode counters (Alpha Science and Ordela) the gas was a combination of purified argon with a small fraction of methane used to quench the ionization multiplication effect. The ionization mode counter (XIA) used purified argon only. A large electric field is applied across the chamber so that charge carriers (Ar+ and e-) created by the passage of an ionizing particle drift to the high voltage electrodes in the chamber to be collected, amplified and counted as discrete events. The counter records the number of detected events (N) and the length of a counting period (T, in hours). Due to the stochastic nature of radioactive decay, the count rate, N/T, has a Poisson distribution with a mean value of N/T. When N is greater than 20, which is true for all measurements in this study, the Poisson distribution is accurately approximated by the normal distribution with mean value N/T and standard deviation equal to $\sqrt{N}/T$. This random variation in count values is referred to as "counting uncertainty" to differentiate it from other systematic errors such as equipment calibration error or sample area measurement.

All alpha counters have a "background" rate caused by alpha emissions from the counting chamber walls, alpha emissions from the sample tray, electronic noise in the detector and other sources. During sample counting, with the sample in the counting chamber, both the sample emission and the background events are present so the resulting count values of N and T are subscripted with "S+B", reflecting the contribution both from the sample and the background. Depending on the counter design the background events may either be discriminated from sample emissions and rejected during counting, or they may be estimated by measuring the count rate when no sample is present and subtracting this value from the count rate when sample and background are both present. Counting values made without a sample in the chamber are subscripted with "B" referring to the background. The corresponding count rates for these cases are simply N_{S+B}/T_{S+B} and N_B/T_B, respectively. All counters also have an associated efficiency (ε), empirically determined during calibration and primarily related to the counter geometry, which expresses the fraction of counts recorded vs. those emitted from the sample. The reader is referred to [8] for a thorough discussion of counters and counting statistics.

II. METHODS

Due to the sensitive nature of these measurements, the study was designed to protect the identity of each participant and only publish values identified by anonymous laboratory codes. It was decided that this lack of transparency does not undermine the primary objective of the study which is to quantify the variability of alpha emissivity measurements as typically employed by manufacturers at all levels of the semiconductor industry. Author C.S. and his company, Ops A La Carte, were contracted to execute confidential disclosure agreements with all participating labs and manage all data reports. All counting reports and the results of the final survey were gathered by C.S. during the experiment. Identifying information was removed, coded identifiers were assigned to each participant and the complete, disguised dataset was provided for analysis at the conclusion of the data collection phase. Some particular patterns that could be inferred from the full data set necessitated that the coded identifiers be assigned independently for the LA samples and for each ULA sample to protect confidentiality. For similar reasons the survey results are only reported in summary rather than being attributed to specific lab codes. The records of the original data are maintained by C.S. for reference. Each participant's lab codes and results were sent to participants individually at the conclusion of the data collection.

Counting samples for this study were prepared representing an ultralow alpha (ULA) material with an emissivity of less than 2 khr^{-1}-cm^{-2} and low alpha (LA) material with an emissivity from 2 to 50 khr^{-1}-cm^{-2}. Due to the limited time available for collecting data and the large number of participants, 2 matched samples of the ULA and LA material were prepared. These samples are referred to as ULA-1, ULA-2, LA-1 and LA-2. Low alpha (LA) samples were prepared by melting 135 g of commercial grade aluminum 6061 alloy (170 khr^{-1}-cm^{-2}) and 435 g of 99.9995% pure aluminum (0.3 khr^{-1}-cm^{-2}) at 700 C in a nitrogen atmosphere. After solidification, all surfaces were machined to remove oxidation or scale and the ingots were rolled to a thickness of 1 to 2 mm, and then sheared into appropriately sized sections to fit the counting chambers. All pieces were cleaned with caustic and acid solutions followed by deionized water rinse and high purity isopropanol wipe. The resulting samples were stored in bags backfilled with nitrogen. Each LA sample was provided as a set of 4 pieces totaling more than 1000 cm^2.

Ultra low alpha (ULA) samples were prepared from 99.995% pure titanium bars that were rolled to 0.5 mm, cut, cleaned and then stored similar to the LA samples. The ULA samples were narrower and required 8 pieces to provide adequate area of more than 1000 cm^2.

To confirm emission stability and a reasonable match between replicates, all 4 samples were counted multiple times over a period of several weeks prior to the start of data collection and again once at the end by B.M.C using a common protocol and equipment. Pre- and post- measurement counting results are provided in table 1. Values for emissivity are in percent, normalized to the mean values of the LA and ULA measurements. This choice was made since there is no available counting standard to provide an absolute calibration

TABLE I: PRE- AND POST- DATA COLLECTION MEASUREMENTS TO ASSESS THE STABILITY AND MATCHING OF SAMPLE EMISSIVITY.

Sample	Relative Emissivity		Time (days)
	Pre-	*Post-*	
LA-1	92 ± 4	93 ± 6	224
LA-2	115 ± 7	101 ± 4	207
ULA-1	150 ± 60	100 ± 80	238
ULA-2	100 ± 60	40 ± 60	239

and these measurements are only to check the stability and match among the samples. The "Time" column records the number of days between the pre- and post- emissivity measurement. Although the values in table 1 might suggest some emissivity mismatch or instability, particularly in the LA samples, the reader is referred to the results reported for all paired LA samples below.

Each sample was labeled with a coded identifier that was changed prior to being sent to each lab for counting, blinding the participant to the particular sample's identity. The label also instructed the participant to count the reverse side of the sample to ensure that only 1 surface was used. Samples were enclosed and heat sealed in a commercial polyethylene bag material before being shipped for counting. Participants were told if the sample they received was LA or ULA so that they could adjust their procedures, if desired (it is common industry practice to adjust the counting time or sample area based on the expected level of activity). No other specification was made with regard to procedure so that the full range of methodological variation would be sampled. At the conclusion of a measurement the participant filled out a data collection form and returned it to C.S. where identifying information was removed. At the end of the data collection phase the full, disguised data set was provided for analysis. A smaller and more fully disguised subset of data was provided to J.D.W. early in the collection phase to confirm the procedures.

By design, the order of sample counting was not randomized, although it was not disclosed to the participants. The objective of the counting order was to have samples LA-1 and ULA-1 sent repeatedly to a subset of participants, while samples LA-2 and ULA-2 were sent to each participant once. Due to time constraints it was not possible to completely meet this schedule. The participant code assignment was a result of patterns discernible in the partial data set and the requirement to protect the participant's confidential information.

TABLE II: EXPERIMENTAL SUMMARY OF THE COLLECTED DATA. EMISSIVITY VALUES ARE IN KHR^{-1}-CM^{-2}.

	LA		ULA	
	1	**2**	**1**	**2**
Number of participants counting this sample	6	9	6	9
Samples submitted to participant	10	9	10	8
Data forms submitted	10	9	13	10
Reported Values				
Mean / median reported mean emissivity	31.2 29.15	30.6 29.8	0.8 0.8	0.7 0.6
Minimum / maximum reported mean emissivity	20.2 45.3	20.3 45.7	0.22 1.6	0.0 0.812
Minimum / maximum reported upper bound	—	—	0.19 0.8	<0.3 1.4
JEDEC Calculated Values				
Consensus mean / median α	31.0 29.2	30.6 29.8	0.7 0.7	0.8 0.1
Minimum / maximum α	20.2 45.1	20.2 45.5	0.0 1.3	0.0 0.8
Minimum / maximum counting uncertainty (1 σ)	1.8% 10. %	1.8% 13. %	20% 67%	56% 210%

Following data collection a voluntary survey was sent to each participant with the instructions that no proprietary information was to be disclosed. Each question included an option for "no response" that was used to indicate these purposeful omissions. Surveys were gathered by the C.S., identifying information was removed and the remaining responses were summarized.

III. RESULTS AND OBSERVATIONS

Table 2 summarizes the number of participants measuring the different samples and their reported values. The JEDEC calculated values will be discussed in detail below but are included here for comparison. Each sample was sent to a participant for counting one or more times and in several cases more than one count was performed by a participant on a particular round. The number of measurement rounds for each sample and the total number of measurements are both detailed.

Table 3 details all values from the data collection in the "Reported Values" columns. Each participant's readings have been assigned a code in the format "SS.PPNa" to identify the sample; participant, order, for paired readings; and participant order, when multiple readings were made by a participant. Sample codes are L1, L2, U1 or U2 for samples LA-1, LA-2, ULA-1 and ULA-2, respectively. The identifying code for LA samples is a single character while ULA samples use a number-letter or letter-number code for ULA-1 and ULA-2, respectively. For participants that made measurements of the same sample at different times during the study, the participant code is followed by "1" or "2," to denote the order that the readings were made. Finally, for participants that provided multiple measurements when they received a sample a lower case alphabetic character is assigned to differentiate them. The order of these multiple readings, other than those for U1.4G1a and U1.4G1b, was not available from the collected data so temporal order should not be assumed.

Data presented in table 3 have been appropriately censored, as defined by each participant's counting procedure, to remove unrepresentative data at the beginning of a counting cycle when contaminants are being flushed from the counting chamber. The "Reported Values" columns are directly as reported on the data forms, with the following exceptions:

- All emissivity and error values have been uniformly reported in units of khr^{-1}-cm^{-2}, although other units may have been used on the data collection forms. The same number of significant digits and any other notations have been retained.
- The reporting of background information was not uniform due to variations in participant procedures. Although the data collection form requested values for N_B and T_B, some reports instead provided N_B/T_B and T_B. In these cases, the value of N_B was inferred from the data. For participants

reporting N_B and T_B the calculated value of N_B/T_B is reported.

- Participants were requested to report emissivity as an estimated value or upper bound, at their discretion. These are denoted with an upper case 'E' or 'B' when reported.
- All inferred or calculated values are shown in red italics.

The precise definition of the reported emissivity and error values are self-defined by the participant, as is the choice of

TABLE III: SUMMARY OF ALL DATA REPORTED.

		REPORTED VALUES										JEDEC			
Lab	Reading	A (cm2)	ε	Ns+b	Ts+b (hrs)	Nb	Tb (hrs)	Nb/Tb	Est/ UB	Emiss-ivity	Error	α	σ	$LOD_{90\%}$	Notes
Sample LA-1 (α_{mean}=31.0 khr^{-1} cm^{-2}, 20.2 < α < 45.1, 1.8% < σ < 10%)															
A	L1.A	804.1	88%	2120	60	548	159	*3.45*	E	45.3	1.1	45.1	1.1	1.8	
B	L1.B1	1000	84%	901	48	*86*	48	1.8	E	20.2	0.8	20.2	0.8	1.3	
B	L1.B2	1000	84%	1024	48	*149*	48	3.1	E	21.7	0.8	21.7	0.8	1.4	
J	L1.J1	706.86	93.96%	597	20.5			0	E	44.945	1.783	43.8	1.8	2.9	
J	L1.J2	706.86	93.84%	157	6			0	E	39.449	3.05	39.4	3.1	5.2	
N	L1.N1	848	90%	1401	48	173	48	*3.60*	*e*	33.5	2.2	33.5	1.1	1.8	
N	L1.N2	848	90%	1413	48	183	48	*3.81*	*e*	33.6	2.2	33.6	1.1	1.8	
Q	L1.Q	866	84%	1947	92	320	92	*3.48*	U	24.4	0.3	24.3	0.7	1.2	
Y	L1.Y1	848	90%	2324	100	431	100	4.31	E	24.8	0.7	24.8	0.7	1.1	
Y	L1.Y2	840	90%	3322	150	602	150	*4.01*	E	24	0.55	24.0	0.6	0.9	
Sample LA-2 (α_{mean}=30.6 khr^{-1} cm^{-2}, 20.2 < α < 45.5, 1.8% < σ < 13%)															
A	L2.A	546.48	88%	4186	166	403	120	*3.36*	E	45.7	0.9	45.5	0.9	1.4	
B	L2.B	1000	84%	1122	48	*149*	48	3.1	E	24.1	0.9	24.1	0.9	1.5	
D	L2.D	754	84%	1715	87	375	87	*4.31*	E	24	0.8	24.3	0.8	1.4	
J	L2.J	706.86	94.24%	110	4			0	E	41.28	3.82	41.3	3.9	6.5	
N	L2.N	852	90%	1452	48	183	48	*3.81*	*e*	34.5	2.2	34.5	1.1	1.8	
Q	L2.Q	548	84%	1199	92	344	92	*3.74*	U	20.3	0.3	20.2	0.9	1.5	
V	L2.V	1065	84.20%	2902	95	*1925*	500	3.85	*e*	29.8	0.8	29.8	0.6	1.0	
Y	L2.Y	848	90%	2248	100	602	150	*4.01*	E	24.2	0.7	24.2	0.7	1.1	
Z	L2.Z	950	85%	4691	164	462	164	*2.82*	*e*	31.9	0.9	31.9	0.5	0.9	
Sample ULA-1 (α_{mean}=0.7 khr^{-1} cm^{-2}, -0.5 < α < 1.3, 20% < σ < 67%)															
1A	U1.1A1	940	84%	427	87	369	87	*4.24*	E	0.8	0.4	0.8	0.4	0.7	
1A	U1.1A2	940	84%	369	87	341	87	*3.92*	E	0.4	0.4	0.4	0.4	0.6	
2E	U1.2E	900	85.60%	444	79	*469*	100	4.69	*e*	1.6	0.4	1.2	0.4	0.7	
4G	U1.4G1a	799.8	90%	417	100	431	100	4.31	U	0.19	0.38	-0.2	0.4	0.7	1
4G	U1.4G1b	799.8	90%	417	100	602	150	*4.01*	E	0.22	0.38	0.2	0.4	0.6	1
4G	U1.4G2	757.8	90%	328	150	123	67	*1.84*	E	0.514	0.24	0.5	0.3	0.5	
5H	U1.5H1	706.86	89%	30	66			0	*e*	0.722	-0.131/-0.275	0.7	0.1	0.2	
5H	U1.5H2	706.86	90.64%	21	40			0	E	0.819	+.170/-0.253	0.8	0.2	0.3	
8U	U1.8U1a	505.8	81%	446	138	948	322	*2.94*	E	0.6	0.4	0.7	0.4	0.7	2
8U	U1.8U1b	528	90%	396	138	803	326	*2.46*	E	1	0.4	0.9	0.4	0.6	2
8U	U1.8U1c	649.8	81%	491	162	803	326	*2.46*	E	1.1	0.3	1.1	0.3	0.5	2
8U	U1.8U2	633	81%	335	108	565	234	*2.41*	E	1.3	0.4	1.3	0.4	0.6	
9F	U1.9F	806.2	84%	363	116	401	116	*3.46*	U	0.8		-0.5	0.4	0.6	
Sample ULA-2 (α_{mean}=0.3 khr^{-1} cm^{-2}, -0.6 < α < 0.8, 56% < σ < 210%)															
A5	U2.A5	825	84%	370	87	409	87	4.7	U	< 1	0.4	-0.6	0.5	0.8	
C6	U2.C6	1000	84%	164	46	*156*	52	3	E	0.7	0.4	0.7	0.4	0.7	
H2	U2.H2	1000	85%	932	332	1006	332	*3.03*	U	< 0.3	0.3	-0.3	0.2	0.3	
L0	U2.L0	818	90%	214	48	223	48	*4.65*	*u*	1.4	1.1	-0.3	0.6	1.0	
M9	U2.M9	819.5	90%	162	70	123	67	*1.84*	E	0.65	0.33	0.6	0.3	0.5	
N4	U2.N4	706.86	92.70%	25	48			0	E	0.812	+.156/-.395	0.8	0.2	0.3	
P8	U2.P8a	537.4	90%	378	113	743	240	*3.10*	E	0.6	0.4	0.5	0.4	0.7	2
P8	U2.P8b	538.6	81%	326	113	692	240	*2.88*	E	0.0	0.4	0.0	0.4	0.7	2
P8	U2.P8c	583.2	90%	404	120	872	265	*3.29*	E	0.1	0.3	0.1	0.4	0.6	2
Q1	U2.Q1	828	84%	276	92	281	92	*3.05*	U	0.8		-0.1	0.4	0.6	

Notes

1. Reading U1.4G1a is reported as an upper bound and includes a note that the background data was taken before sample counting for 100 hours. Reading U1.4G1b is reported as an estimated value and background data was taken after counting for 150 hours. The count data is identical for these measurements.
2. Notes on the data collection form indicate that various combinations of sample pieces and counters were used for these measurements.

reporting an estimated value or an upper bound. Three columns have been added to right hand side of table 3 for comparison with these data. The columns labeled "α" and "σ" are calculated from the counting data (N_{S+B}, N_B) and parameters (T_{S+B}, T_B, A and ε) according to the definition in the proposed JEDEC counting standard [6]. These values are the expected value for the emissivity, in khr^{-1}-cm^{-2}, and the counting uncertainty expressed as the standard deviation in the same units. The "$LOD_{90\%}$" column, also from the proposed JEDEC standard, is the "limit of detection" expressed at the 90% confidence level. This value is the minimum emissivity required from a sample so that it can be discriminated from the background at the 90% confidence level. The $LOD_{90\%}$ is determined by the choice of counting parameters and is proportional to the counting uncertainty. Specifically, the definitions from the JEDEC proposal are:

$$\alpha = \frac{\frac{N_{S+B}}{T_{S+B}} - \frac{N_B}{T_B}}{A\,\varepsilon} \times 1000, \qquad (1)$$

$$\sigma = \frac{\sqrt{\frac{N_{S+B}}{T_{S+B}^{\,2}} + \frac{N_B}{T_B^{\,2}}}}{A\,\varepsilon} \times 1000, \quad \text{and} \qquad (2)$$

$$LOD_{90\%} = 1.64\,\sigma. \qquad (3)$$

A. *Low alpha measurements*

Figure 1 presents the emissivity measurements for the LA samples, grouped by participant.

It is immediately apparent that each lab provides consistent results for its own measurements, independent of the sample or measurement order. None of the 19 measurements are within 1 σ of the consensus mean value of 30.9 khr^{-1}-cm^{-2}, indicated by the purple dashed line, and only 2 (L2.V and L2.Z) are within 2 σ.

Participants B, J, N and Y each measured sample LA-1 twice with an average of 69 days between readings. When considered pair-wise for the first vs. the second measurement (e.g., L1.B1 vs. L1.B2) the average deviation between samples is -0.2 σ, with a range of -1.8 to +1.2 σ, supporting the assumption that sample LA-1's emissivity is stable over time these measurements were made. These readings may also be combined with the pre- and post- data collection readings made by B.M.C (table 1) for 5 paired measurements. In this case the resulting average deviation is 0.0 σ and the range remains the same, again supporting an assumption of stability.

In a similar analysis participants B, J, Q, N and Y made measurements of both LA-1 and LA-2. By averaging the multiple measurements of LA-1 made by B, J, N and Y a set of 5 pair-wise comparisons may be made for LA-1 vs. LA-2. The

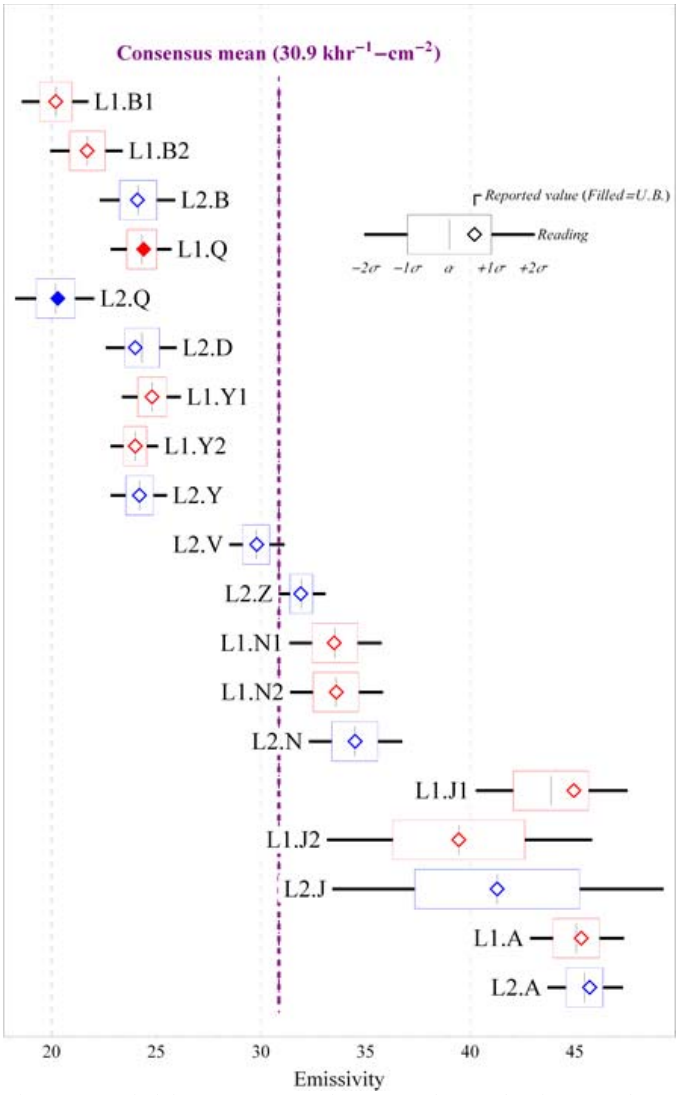

Figure 1. Emissivity measurements for LA-1 and LA-2 showing more than a 2X range in the estimated values. The key for interpreting each measurement is shown in the figure. Filled diamonds represent values reported as upper bounds and empty diamonds are reports as estimated values.

average deviation for these readings is 0.0 σ with a range from -4.4 σ to +3.6 σ, supporting the assumption that sample LA-1 and LA-2 are matched. These measurements may also be augmented with the readings from table 1 resulting in an average deviation of 0.1 σ and no change in the range.

B. *Ultralow alpha measurements*

The large counting uncertainty and mixture of upper bounds vs. estimated values in these reported results makes the reported emissivities very difficult to compare directly. All of the values are available in table 3 for the interested reader and some commentary is provided in the discussion in this regard. The remainder of the quantitative discussion therefore focuses on the calculated values for α and σ determined from the more objective reports of counts and time.

Figure 2 presents the 13 measured values for emissivity (α) from 6 participants for ULA-1 and Fig. 3 presents 10 measurements from 8 participants for ULA-2. The purple dashed line indicates the consensus mean values of 0.6 and 0.2 khr^{-1}-cm^{-2} for ULA-1 and ULA-2, respectively. Paired measurements of sample ULA-1 were made by participants 1A, 4G, 5H and 8U, separated by an average of 93 days. Participants 2E and 9F each contributed a single reading for ULA-1. Eight of the 9 participants were able to provide measurements for ULA-2.

Six of 13 measurements for ULA-1 and 5 of 10 for ULA-2 include the consensus mean within the 1σ range of their counting uncertainty. This distribution of measurements does not demonstrate the range of emissivity values seen in the LA data, where the counting uncertainty is so much smaller than the range of emissivity measurements that it is obvious that they are distinctly different values. It may be possible that similar lab dependence exists in these measurements but is masked by the counting uncertainty. Unfortunately, the large uncertainties lead to qualitatively weak statistical conclusions that there is no evidence of a difference.

Participants 4G, 5H, 1A and 8U each measured sample ULA-1 twice with an average of 93 days between readings. Considered pair-wise, the average deviation between samples is +.4 σ, with a range from -1.1 σ to 1.2 σ. This supports the assumption that sample ULA-1's emissivity is stable over time.

C. *Participant survey*

The results from the post-data collection survey are summarized in table 4. One participant did not reveal the counter that was used, 1 used an Ordela 8600A-LB, 1 used an XIA UltraLo-1800 and the remaining 6 were Alpha Sciences Model 1950. Six of 8 labs followed a written procedure for counting, and only 1 of 9 reports taking any "extra" care with this these measurements compared to others that are routinely made. Determination of a background level was done in several ways combining measurements made immediately before or following a sample count with historical data. Procedural choices, such as pre-cleaning of samples prior to counting, handling precautions, and packaging for return shipment varied widely. The energy detection threshold was also a variable, ranging from 1 to 3 MeV for reported values. The impact of this particular variation is discussed in more depth below. It is hoped that the survey results will be helpful for formulating hypotheses for future studies in this area.

IV. Discussion

A. *Variability in LA data*

The most striking observation that can be made from these data is the variation of approximately 2X in the reported emissivity values for the LA samples, correlated with the lab making the measurement. From the survey results it is possible to estimate the magnitude of the impact of errors in the measurement of the sample area, counter efficiency and background determination, relative to the expected emissivity, as being well under 10% each. The resulting combined error is therefore expected to be less than 20% from these random factors. Eight of 9 counters were calibrated within 24 months making this an unlikely source of additional error. Although there are many unexplored sources for potential error from

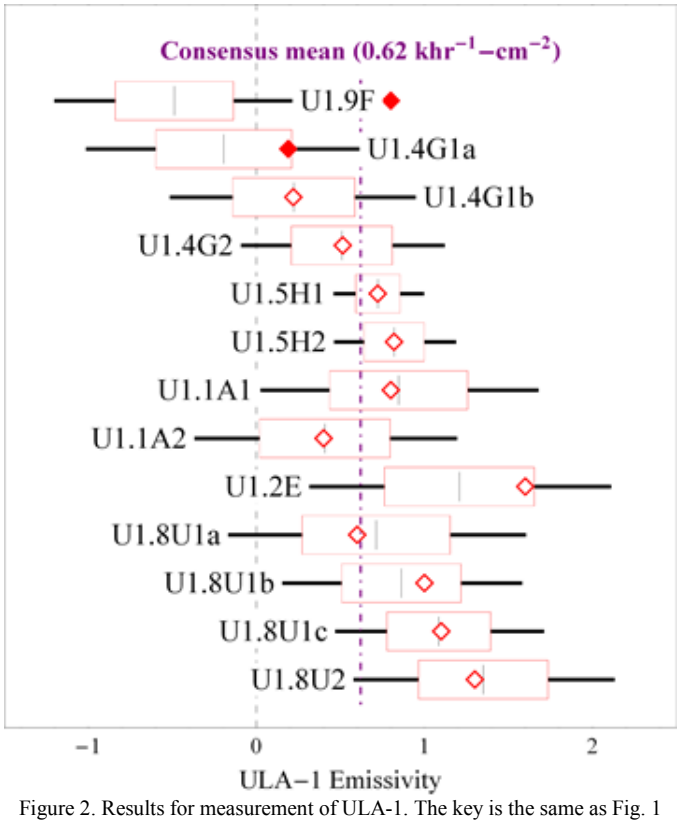

Figure 2. Results for measurement of ULA-1. The key is the same as Fig. 1

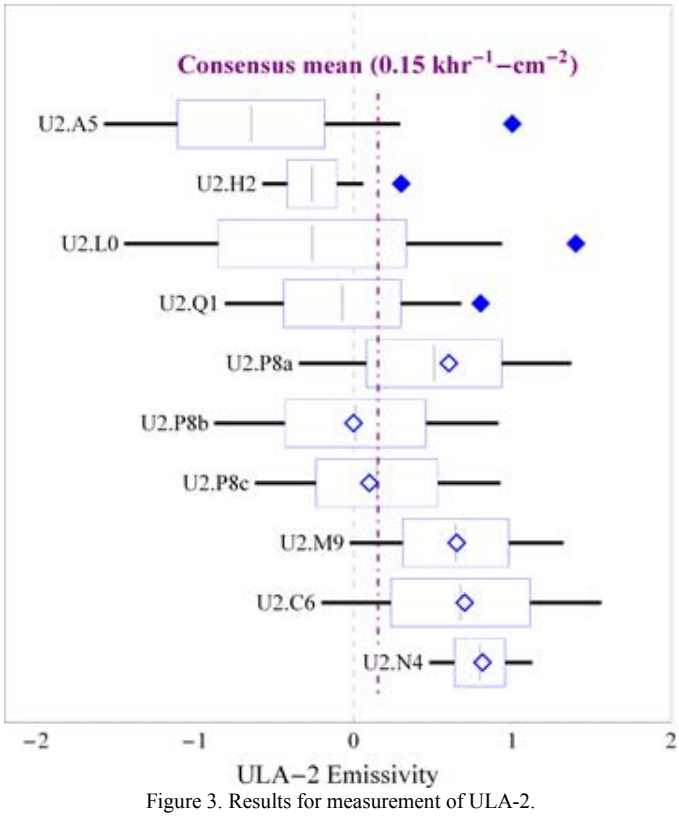

Figure 3. Results for measurement of ULA-2.

procedural differences, such as sample contamination, improper procedures, and misreporting it is the authors' opinion that these would not explain the pattern seen in these data.

It is likely that most of this measurement difference is not the result of counting uncertainty or measurement process capability. Given the strong correlation of the variation with a particular lab that persists when the JEDEC values are calculated using a single procedure, it is quite likely that the variation is linked to the particular counter used at each site.

TABLE IV. SUMMARIZED RESULTS FROM PARTICIPANT SURVEY.

	Question	Responses
1	Was a written procedure followed for this testing?	*NR* - 2,*Yes* - 6, *No* - 1
2	Was the testing for this experiment conducted by personnel that are generally responsible for counting operations?	*Yes* - 9, *No* - 0
3	Was any "extra" care taken with this experiment? For example, were additional tests run that are not normally required, extra quality checks conducted, counters used other than those normally used for production?	*Yes* - 1, *No* - 8
4	What brand(s) and model(s) of counter were used for LA samples?	*NR* - 1, Ordela 8600A-LB - 1, Alpha Sciences 1950 - 6, XIA Ultralo 1800 - 1
5	What brand(s) and model(s) of counter were used for ULA samples?	*NR* - 1, Ordela 8600A-LB - 1, Alpha Sciences 1950 - 6, XIA Ultralo 1800 - 1
6	Was the counter for this experiment built in-house or modified in-house?	*NR* - 5, *Built in-house* - 3, *Modified in-house* - 1
7	What is the lower detection threshold for the counter used for LA samples?	*NR* - 3, 1 MeV - 2, 1.1 MeV (Est.) - 1, 3 MeV - 1, *Other* [no description] - 1 , Other - LLD of 0.9 khr-1 cm-2
8	What is the lower detection threshold for the counter used for ULA samples?	NR - 3, 1 MeV - 2, 1.1 MeV (Est.) - 1, 3 MeV - 1, *Other* [no description] - 1, Other - LLD of 0.8 khr-1 cm-2
9	How was the background determined for LA samples?	*NR* - 1, *Measured and combined with historical data* - 1, *Measured after counting* - 0, *Measured before counting* - 4, *Historical data only* - 1 , Measured every month before counting - 1, Background measured earlier in the month - 1 (notes2, 3, 4)
10	How was the background determined for ULA samples?	*NR* - 1, *Measured and combined with historical data* - 1, *Measured after counting* - 0, *Measured before counting* - 4, *Historical data only* - 1, Measured every month before counting - 1, Background measured before and after sample] - 1 (notes 2, 3 & 4)
11	How was the length of the background count determined for LA samples (for either per-sample or historical measurements)?	*Same as the length of the sample count* - 4, *Standard length of time based on sample type* - 1, *Standard length of time, independent of sample type* - 2, No background - 1 (note 5)
12	How was the length of the background count determined for ULA samples (for either per-sample or historical measurements)?	*Same as the length of the sample count* - 4, *Standard length of time based on sample type* - 0, *Standard length of time, independent of sample type* - 2, No background - 1 (note 5), *Other* (note 5)
13	What is in the counting chamber during background determination?	*NR* - 1, *Only the sample tray* - 3, *Known emissivity standard* - 0, Zero or near-zero emissivity standard - 1, *Other* (note 5), "Empty" - 1
14	Was a correction made to the background to account for the portion of the sample tray that would be covered by the counting sample?	*NR* - 2,*Yes* - 0, *No* - 7
15	Does the counter efficiency contribute to the background determination?	*NR* - 2, *Yes* - 7, *No* - 0
16	Were samples handled according to your normal procedure or work flow?	*Yes* - 9, *No* - 0
17	Was the sample cleaned or otherwise prepared for counting? If so, how?	*Yes* (please describe) - 4 (notes7, 8, 9), *No* - 5
18	Was the sample packaged for return according to a standard procedure? If so, what packaging was used?	*NR* - 2, *Yes* (packaging in which sample was received) - 4, Yes ("standard bag used for shipping low alpha material") - 1, *No* - 2
19	When was the counter last calibrated?	*NR* - 1, < *6 months* - 3, < *12 months* - 2, < *24 months* - 2, *Other (please describe)* - 1 (note10)
20	Who performs counter calibration?	*Manufacturer* - 4, *Independent laboratory* - 0, *In house* - 5
21	How was counter efficiency determined?	*Manufacturer* - 7, *Independent laboratory* - 0, *In house* - 2
22	What criteria were used for determining the counting time?	*NR* - 1, *Single time used for all samples* - 0, *Based on early count results* - 0, *Based on expected sample activity* - 5, *Other (please describe)* - 3 (notes 11, 12 & [1]3)
23	What criteria were used for determining the area of the sample to be used for counting?	Maximized area able to fit in counter - 7, Standard area for all samples - 1, Other (please describe) - 1 (note 14)
24	What is your estimated error of the sample area measurement?	*NR* - 2, < *1%* - 1, < *2%* - 2, < *5%* - 1, < *10%* - 1, *Other (please describe)* - 2 (note [1]5)

Responses in italic type are from the survey multiple choices. Non-italic responses are provided by the respondent. Longer, non-standard, responses are provided as notes. NR indicates "No Response"

2. One respondent that measured background before counting noted that it may alternatively be measured following counting if the schedule is tight.
3. "Sample areas are sufficient to cover region which produces background from the sample tray. Other sources of background are included in the count and can't be measured independently yet."
4. Only one respondent had different background determination procedures for LA vs. ULA samples.
5. One response (same respondent as note 3) said that there is 0 background due to the counter design.
6. "Other, Time to achieve reasonable relative Standard deviation but maximize sample throughput."
7. "Yes, ..." Solid samples are cleaned with ethanol wipe and then vacuumed. Powder and polymer samples are spread on an acrylic plate, covered with Mylar film and stored in flowing P-10 for at least 36 hours.
8. "Yes, Sample was not cleaned, but arranged on tray in a N2-filled glovebox."
9. "Yes, Wiped with high purity isopropanol."
10. "Other, Thorough calibration 13 months ago, calibration check 3 months ago."
11. "Other, Comb[ination] of expected activity and early results, rounded to nearest convienent time.
12. "Other, Sample emissivity relative to background."
13. "Other, For solid sample: 48hrs. For colloid or powder sample:72hrs.
14. "Other, Since the counting area of the sample is set by the electrode size, and the electrode size is a manufactured part, the error in the electrode size is very small." (same respondent as notes 3 and 5)
15. "Other, Hard to estimate the error with one number. Some samples were irregularly shaped, some were not."

One mechanism that could explain the variation is hinted at by the response to survey questions Q7 and Q8, reporting a range of energy thresholds used by each participant. All of the alpha counters in this study use some sort of energy threshold to differentiate between alpha and beta particle emission, and to prevent electronic noise from introducing spurious counts. Alpha particles entering the count chamber at low energies are therefore eliminated from the counting.

For solid materials the maximum range of an alpha particle is less than 100 μm. Materials with alpha emissions spread to a depth of more than a few 10's of micrometers are colloquially referred to as "thick" while materials with the emissions concentrated very near the surface are "thin." Thick samples are most commonly encountered and are representative of emissions from bulk source materials. Thin samples are less common and generally associated with surface contaminant deposition. This study used thick samples for both the LA and ULA material.

For these thick samples it is assumed that the alpha particle energies entering the sample chamber are distributed from 0 up to several MeV. The reader is referred to [4] where a simulation of a thick sample's alpha emission energies are shown in figure 2. The counter's threshold setting determines the fraction of these alpha particles that are rejected during counting. As the particular isotopes and their relative ratios are unknown for these samples it is not possible to make a quantitative estimate of this effect. It is also not possible to compare the energy threshold of each site with the reported values using this data set due to restrictions related to confidentiality. All that can be said is that this mechanism would be expected to create a variation between counts made using different equipment, while demonstrating a strong repeatability for measurements made on the same equipment. This observation is consistent with the data. A more rigorous test of this hypothesis awaits the availability of a different data set.

Although a precise statistical test has not been developed yet, it is believed that the weak trends in the ULA data are not inconsistent with the more robust findings for the LA data.

B. Inconsistent reporting conventions

Comparing the reported emission values from each lab with the mean and standard deviation calculated according to the JEDEC method reveals differences, particularly in the ULA data. It is not expected that the participant's reported value would necessarily match the JEDEC value, due to different assumptions regarding the counting process capability, interpretation of the reported values and their intended application. Nevertheless, the JEDEC values of α (sample mean) and σ (standard deviation) are useful metrics to highlight the choices made in this reporting.

As an example, measurements of ULA materials are commonly reported as an upper bound due to the high level of uncertainty associated with acceptable choices for counting parameters. For these reports, it is reasonable to assume that the emissivity and the portion of the error from counting uncertainty would be reported as

$$\begin{aligned} Emissivity_{U.B.} &= \alpha + N\sigma \\ Error_{U.B.} &= \sigma \end{aligned} \tag{4}$$

The value of N would be chosen to reflect the desired confidence level of the reported result and would be expected to range from 1 to 3, corresponding to confidence levels of 84% and 99.9%, respectively. A variety of choices are reported in table 3 with values of N ranging from 1 to 2.3. Reading U1.9F reports an emissivity equal to 2σ and an unspecified error. Two LA readings, L1.Q and L2.Q, are reported as upper bounds although the emissivity is equal to α in both cases.

There is no "correct" method regarding this reporting. The observed variability in these values emphasizes the need for a standard reporting method, a clear definition of the meaning of the numerical values being reported, or both. The JEDEC proposed standard requires that the mean emissivity and the standard deviation of the emissivity both be reported, along with many other parameters that can be helpful in interpreting the report.

C. Sample quality

Referring to the paired comparisons presented earlier, these data do not demonstrate a mismatch in emissivity between LA-1 and LA-2, or instability in emissivity for LA-1 or ULA-1. By association, it is reasonable to assume both that LA-2 and ULA-2 are also stable given their common starting source materials and subsequent processing (i.e. LA-2 matched with LA-1 and ULA-2 matched with ULA-1). The only data available to compare the matching of ULA-1 and ULA-2 is the single paired measurement made by B.M.C. before and after the data collection phase (see table 1). Pooling the before and after data, under the assumption of stability, the difference in the mean value of these measurements is 0.5 σ, consistent with an assumption that the samples are matched. Given the large uncertainty in the sample measurements this conclusion is necessarily weak statistically. It is not unlikely that the ULA and LA samples are both matched and stable. Given the apparent match between samples it is very likely that the activities of the individual pieces making up each sample are also well matched. There is a possibility, that is not explored here, that the emissivity of the individual pieces could vary due to differences in surface contamination but this is considered unlikely.

D. Role of Counting Uncertainty and Background Determination

When making measurements of emissivity very near the limit of the instrument's capability, the accurate determination of the sample+background and background counting rates is extremely important. From the measurements shown in table 3 it is evident that the range of N_B/T_B (1.84 to 4.70 hr-1) substantially overlaps the range of N_{S+B}/T_{S+B} (0.45 to 5.62 hr-1) values. Three example ULA measurements are shown in fig. 4. For each measurement the values for N_{S+B}/T_{S+B} and N_B/T_B are shown as individual points accompanied by the distribution of the counting uncertainty. The distance between these distributions, shown by the dotted lines, is the net count rate, in counts per hour, from which the emissivity is calculated.

Measurement U1.8U2 demonstrates adequate separation between the rates, while the differences for measurements U2.C6 and U2.L0 are so small, compared to the counting uncertainties that the net difference is overwhelmed by the uncertainty. The lower right corner of Fig. 4 simulates the effect of increasing the count times by 4X for measurement U2.C6, reducing the uncertainty to a level where the difference can be adequately determined.

Small errors in the background determination, either due to counting uncertainty or systematic error, can easily result in a substantial error in the measured emissivity. Likewise, inadequate counting time results in large uncertainty that will overwhelm the small difference in rates that are trying to be discerned. The proper choice of counting parameters and the careful control of the background rate are obviously vitally important when making measurements for ULA samples.

V. Conclusions

These preliminary data demonstrate that there is a wide variation for LA emissivity measured by various centers that cannot be attributed to normal statistical variations inherent in the measurement that are typically referred to as "counting uncertainty." Although the ULA data match within the expected uncertainty of the measurements this is more a reflection of magnitude of the uncertainty than confidence in the results. The generally good match for paired measurements made within a lab, and the much greater variability between labs, tends to support this interpretation.

These data should not necessarily be interpreted as a serious problem for the semiconductor industry. Monitoring of a manufacturing process is generally an ongoing operation with all measurements made by a single lab following its own procedure and with a consistent relative calibration method. When interpreted for this application these data support the view that monitoring methods are well behaved, at least for the period of time studied.

Estimations of soft error rates commonly make use of accelerated data taken using a high emissivity source and then scaled by the ratio of the high emissivity source and the LA or ULA target material. Since this ratio is generally greater than 10^8 the determination of each emissivity must necessarily be made using different methods. There are generally a multitude of other uncertainties and assumptions that also have to be quantified during this estimation process. These data help to quantify the uncertainty in this portion of the analysis when scaling from accelerated testing to "as built" devices. Since common practice is for this scaling to be done conservatively it is not likely that the uncertainty will generally invalidate prior analyses. Still, it is prudent to consider these findings and assess their impact on soft error estimates.

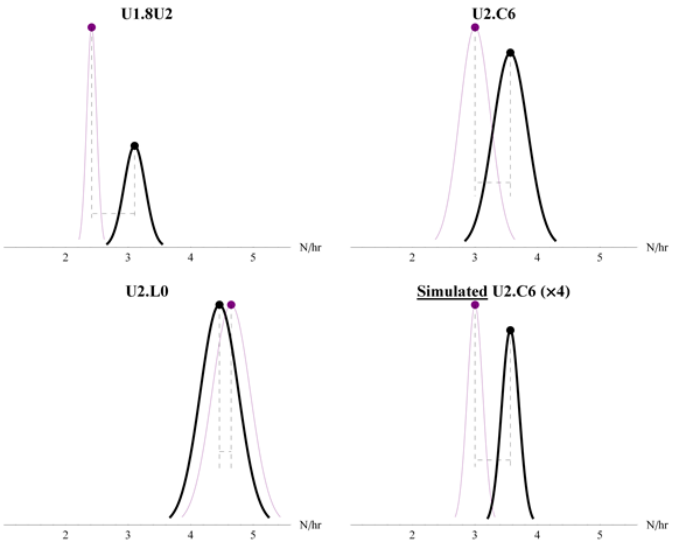

Figure 4. Sample+background and background rates shown with the range of uncertainty. A simulation of the effect of increasing the values of T_{S+B} and T_B by 4X each for measurement U2.C6 is shown in the lower right corner.

The lack of an accepted emissivity calibration standard in the measurement range of interest may be a significant contributing factor to the LA mismatch. The authors would like to suggest that the development and acceptance of a standard would be a valuable contribution to both the science and application of alpha counting.

After discussing these results the authors are planning additional projects to understand the sources of the variability. We are planning a follow on experiment in which more of the experimental variables are fixed. For example we may work with a single sample, standardize our handling procedures, or require a specific measurement period. It is our hope that further data will quantify the impact of these variables and improve the certainty of these measurements.

Acknowledgements

A project of this scope requires help from a large team of people and the authors would like to express their particular gratitude to Xiaoting Niu, Cisco Systems, who provided early project management and logistical support; Molly Lu, Advanced Semiconductor Engineering, Shi-Jie Wen, Cisco Systems, and Norbert Seifert, Intel, for coordination and encouragement; Roger Kuo, Advanced Semiconductor Engineering, for technical support; Professor J.L. Autran, University of Provence, France, for facilitating measurements in 2 laboratories in the French Alps; Mike Silverman, Ops A La Carte, for managing the confidentiality arrangements; and William K. Warburton, XIA LLC for early guidance and development of the alpha consortium concept. All of the authors also gratefully acknowledge the generous donation of over 9000 hours of valuable counter time by the counting participants, without which this study would not have been feasible. Finally, J.D.W. would like to particularly thank C.S. for going above and beyond his contracted responsibilities by facilitating confidential communications with participants in order to resolve errors, providing thoughtful review and input on this paper, and for identifying every obscure case that could compromise our participant's confidentiality. All of the authors are deeply indebted to him for the success of this project.

References

[1] Baumann, R. C. and D. Radaelli (2007). "Determination of Geometry and Absorption Effects and Their Impact on the Accuracy of Alpha

Particle Soft Error Rate Extrapolations." Nuclear Science, IEEE Transactions on 54(6): 2141-2148.

[2] Kobayashi, H., N. Kawamoto, et al. (2009). Alpha Particle and Neutron-Induced Soft Error Rates and Scaling Trends in Sram. Reliability Physics Symposium, 2009 IEEE International.

[3] Clark, B. M., M. W. Weiser, et al. (2004). "Alpha Radiation Sources in Low Alpha Materials and Implications for Low Alpha Materials Refinement." Thin Solid Films: 384-386.

[4] Wrobel, F., F. Saigne, et al. (2009). "Radioactive Nuclei Induced Soft Errors at Ground Level." Nuclear Science, IEEE Transactions on 56(6): 3437-3441.

[5] Measurement and Reporting of Alpha Particle and Terrestrial Cosmic Ray-Induced Soft Errors in Semiconductor Devices. JEDEC Standard. Arlington, VA, JEDEC: 94.

[6] Clark, B. M. (2010). Personal communication.

[7] Mistry, A., S. Lee, et al. (2000). Characterization of Low Alpha Emissivity System on Electroplated Solder Bumps. 2000 Electronic Components and Technology Conference.

[8] Knoll, G. F. (2010). Radiation Detection and Measurement, John Wiley & Sons, Inc.