Uncertainty Budget in Microwave High-Power Testing

Martín García-Patrón, Manuel Rodríguez, Jorge A. Ruiz-Cruz, *Senior Member, IEEE*, and Isabel Montero

Abstract—Space-borne radio frequency (RF) systems must cope with hard qualification procedures, including the evaluation of high-power handling capability of equipment for space applications. Whatever the electrical parameter is being measured, the general rule of thumb throughout a verification process is to check whether the system can operate up to certain thresholds, which are defined to ensure total reliability for the mission along its operative lifetime. Therefore, assessing and reducing the uncertainty linked to their measurement are mandatory issues as it directly affects the accuracy of the qualification process and hence the safety of the whole space mission. This article presents a novel comprehensive study of all variables affecting measurement uncertainty for high RF power test activities. This study is focused on space applications, and, in particular, multipactor testing, because they comprise the largest number of variables. This is not a restricting case; in fact, the outcome of this work is applicable both for space and ground RF applications. As a conclusion, a complete uncertainty for RF high-power testing is obtained, and, where possible, mitigation actions have also been defined.

Index Terms—Measurement uncertainty, metrology, multipactor, radio frequency (RF) high-power testing, space technology, test security margin, uncertainty budget.

I. Introduction

SPACE engineering has to meet extreme requirements for satellites. These demands are consistent with the huge economic costs of space missions and the fact that any repairing activity is nearly impossible. Therefore, space systems have to succeed in coping with very hard qualification test programs [1]. Through this verification process, the ability of the space system to operate under extreme conditions is demonstrated. The test matrix of a space verification process includes a wide variety of technical disciplines: thermal and vacuum performance, vibration, shock, audible noise, electromagnetic compatibility, electrical parameters, and radio frequency (RF) breakdown phenomena, such as multipactor, corona and power handling, and passive intermodulation (PIM) [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15].

Multipactor breakdown is an electron discharge occurring in equipment operating under high-power RF fields and high-vacuum conditions, such as high-power microwave generators, RF windows, accelerator structures, space-born communication systems, and also large particle accelerators [16]. Multipactor occurs when free electrons accelerated with enough energy by RF fields impact a surface and release secondary electrons. If these secondary electrons can be accelerated and impact the surface again, the number of secondary electrons will grow exponentially therefore disrupting the device operation or even causing system damage. The accuracy of the multipactor threshold is highly determined by the uncertainty of secondary emission properties of the material. In fact, any surface contamination and air exposure or aging can cause an increase in the secondary electron yield (SEY) of the materials [17], [18]. The influence of uncertainty in SEY parameters on the multipactor threshold of rough surfaces was also investigated [19]. To determine the multipactor threshold in RF devices, numerous experimental tests and numerical simulations have been performed all over the world [20], [21]. However, no particular study on the observed uncertainty in experimentally detected multipactor thresholds has been found in the literature to our knowledge.

This work aims to contribute an innovative perspective to the existing technical literature in this topic by providing a systematic, clearer, and more complete uncertainty budget in microwave high-power testing. It is hoped that the renewed proposed approach can help in identifying potential issues and optimizing system performance. By addressing uncertainties and striving for more precise and reliable results, it is believed that this innovation could enhance test efficiency, which could be valuable for the telecommunication industry and scientific research. In addition, this can help improve the design and production processes of RF communication devices and systems for space. As a practical example of this innovation, there has been a clear consensus within the national metrology institutes that this subject required the unified and systematic assessment proposed in this work, addressing the needs pointed out in the framework of a project for the European Agency of Metrology (EURAMET).

Whatever the parameter under test is to be measured, the general rule during the qualification process is to check if

the system can operate up to a certain threshold. This critical level usually includes the so-called security margin, which is an extension over the nominal level for the parameter under test. By doing so, any deviation that may take place during the mission either on the system performance or on environmental conditions can be checked in advance and hence dealt with during the design and manufacturing phases of the project.

Thresholds of parameters, as well as security margins, are usually specified by the standards for use in all European space activities of the European Cooperation for Space Standardization (ECSS). This lead standardization body is a collaboration between the European space industry represented by Eurospace and the European Space Agency (ESA) [22], [23]. Security margins are a cornerstone for space industry and a very sensitive topic: the higher the margins, the safer the mission, but also the design and manufacturing process can be more expensive [24], [25]. On the contrary, a lower margin implies less expensive processes but less reliable results. Here is where metrology shows up as a key factor. Whatever the physical measurand [26], [27], [28], it is necessary to reduce the uncertainty associated with the measurement as it directly affects the security margin being applied. Narrowing down the uncertainty budget for a particular measurement is inversely proportional to the complexity (or cost) of the setup necessary to carry out the measurements. For example, RF qualification tests require very complex setups: many different parts and equipment utilized, extreme environmental conditions ($-160\,°C$ to $+160\,°C$ temperature range, and pressure $<10^{-6}$ hPa), very high-power levels, complex calibration procedures, and so on.

The real-time uncertainty methodology was developed at Keysight to compute material parameters at microwave and millimeter-wave frequencies [29]. In microwave high-power radar test system, the mismatch uncertainty and the attenuation device uncertainty are typically the largest components in the total uncertainty [30]. As far as the authors are concerned, a specific assessment of measurement uncertainty in relation to microwave high-power breakdown testing has not been made yet.

By accurately defining the uncertainty budget and reducing it where possible, the mission security is not compromised, but the cost of design, manufacturing, and especially testing can be significantly decreased. In this scenario, this article is devoted to obtaining a complete uncertainty budget for RF high-power testing activities and to assess whether it may have impact on test security.

Besides, despite the fact that this article is focused on space testing, its rationale and outcome can also be applied to other RF high-power applications: for instance, the development of anti-multipactor surfaces, which also requires very precise insertion loss (IL) and return loss (RL) measurements, the PIM performance of ground stations for 4G and 5G services, the EMC susceptibility testing, and so on [31], [32], [33].

II. INFLUENCE OF UNCERTAINTY IN THE POWER MEASUREMENT

Space-borne equipment characterization regarding RF high-power includes the former mentioned four testing disciplines:

multipactor [4], [5], [6], [7], [8], corona [3], [8], [9], [10], [11], power handling [23], and PIM [12], [13], [14], [15]. In all of them, the device under test (DUT) is characterized through the measurement of its three RF power components: incident power on the DUT (P_i), reflected power by the DUT (P_r), and transmitted power out of the DUT (P_t). Therefore, the identification and accurate estimation of all sources of power measurement uncertainty affect the whole RF validation process.

A. Multipactor Test

Multipactor testing activity has been the chosen framework in this article as it opens the widest range of variables to work with. The multipactor effect is a weak discharge phenomenon threatening RF systems on board of spacecraft [4], [5], [6], [7], [8], [16], [25]. Multipactor occurs under high vacuum conditions (pressure $< 10^{-5}$ mbar) when free electrons, captured by the associated RF electric field, are accelerated and impact on the internal surface of the equipment with sufficient energy to induce the emission of secondary electrons. As the direction of the electric field reversed, these secondary electrons are accelerated again, impact on the opposite surface, and consequently liberate new secondary electrons. As the process is repeated (millions of times per second), an avalanche of electrons rapidly grows, initiating a so-called multipactor discharge. The immediate consequence of this process is the degradation of the RF signal, sometimes surface erosion (if the discharge evolves to corona) and eventually, the failure of the whole transmission path. The most important outcome of a multipactor test is the multipactor threshold, which is the minimum level of incident power ($P_{i,\min}$) at which the breakdown takes place.

A multipactor test is not only related to the discharge threshold but also useful to demonstrating whether the security margin that has been defined will actually guarantee the reliability of the system in real operation. Margins are defined by the ECSS and have been recently updated [22]. The difference between the theoretical multipactor threshold and the nominal power is called the analysis margin. Depending on this margin and on some other parameters (RL, IL, outgassing, etc.), it is decided whether the device requires a test to be definitely qualified. If this is the case, the test must be conducted at nominal power plus a security margin. As it was pointed out earlier, to defining test qualification margins is a sensitive topic: with higher margins, the mission is safer, but also the cost of design and manufacturing is increased. On the contrary, a lower margin implies less expensive processes but less reliable results. In general, the space industry and general-purpose RF manufacturers are interested in the adequate reduction of security margins and in the use of the best RF testing standards, including the uncertainty issue [17], [24]. The uncertainty associated with measurements does play a role in the desirable reduction of margins. In a general sense, with lower uncertainty, the security margins can be reduced.

When designing space-borne RF equipment, a multipactor susceptibility study is carried out by means of analysis techniques, heritage, comparison with similar equipment, etc. Accordingly, a theoretical multipactor threshold is obtained.

Equipment or component type	P1 (dB)	P2 or P3 (dB)	
Heritage	All	Bm[(1)]	Bm[(2)], Cm, Dm
Qualification Test	6	6	6
Batch Acceptance Test	4	4	5
Unit Acceptance Test	3	3	4
[(1)] Heritage by test. [(2)] Heritage by analysis.			

Note: P1, P2 and P3 are categories related to the kind of geometry and material present at DUT's multipactor critical area [22]. Bm, Cm and Dm are DUT types with regards to multipactor performance heritage.
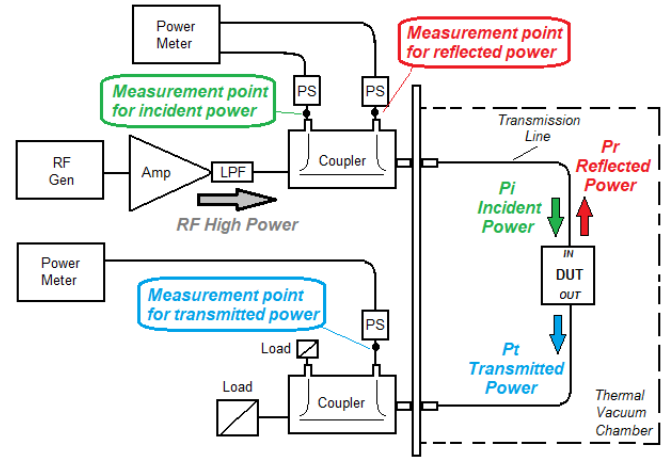

Fig. 1. Typical RF high-power testbed with the DUT located inside a thermal vacuum chamber (TVAC). Critical parameters are the incident, reflected, and transmitted powers of the DUT.

This value is then compared to the nominal RF power the device is envisaged to deal with along its operational life. The amplitude of this margin is dependent on the type of component, as shown in Table I for the so-called P_1, P_2, and P_3 RF equipment or components verified by test [22].

Since multipactor is a physical effect depending on the RF peak voltage, and ultimately, on RF power [2], [4], [5], [6], the key magnitude to be measured on a multipactor test is the multipactor threshold, that is, the RF power level the device is withstanding without experiencing any discharge, or, if it is the case, the RF power level at which the breakdown has eventually taken place. Therefore, the uncertainty associated with measuring RF high power may potentially affect those margins defined in Table I, thus leading to potential noncompliances.

B. Test Margin Versus Uncertainty

The European ECSS-Multipactor Design and Test Standard (ECSS-E-ST-20-01C) [22] defines the multipactor test margin as the required margin of the nominal power with respect to the theoretical multipactor power threshold resulting from an analysis. On the other hand, the batch acceptance margin is defined as the allowance of power over the nominal operational power, during the equipment lifetime, excluding testing, to be applied to any equipment of the same batch. The Aerospace Standard/Handbook for Radio Frequency Breakdown Prevention in Spacecraft Components [2] aims to "minimize potential risks in applicable RF systems and components" by means of a "new and alternative approach (that) removes excessive, hidden, or stacked margins." The RF power measurement uncertainty is another relevant issue that is discussed for the first time in this article. In either case, the RF high-power measurement uncertainty is usually negligible compared to 3 dB (the lowest test margin in Table I).

In the RF literature, it is noticed that some manufacturers and research teams establish some uncertainty tolerance for RF high-power measurements [18], [19]. In addition, it has been published that their "measurement errors in power meters and in test setup calibration are typically less than 3%" [12]. The ESA working paper number 1556 [7], probably the cornerstone in European multipactor research, gives a 6-dB margin

justification (allowing for VSWR, oxidization, contamination, and migration of contamination), and besides states ±1 dB as "measurement error of the test equipment."

In this article, the authors will try to establish a unified uncertainty budget for RF high-power testing and will assess its influence on testing security margins.

C. Uncertainty Sources in RF High-Power Test Setups

The characterization of a DUT relies upon the measurement of the already previously mentioned RF incident power (P_i), RF reflected power (P_r), and RF transmitted power (P_t) of the DUT. Fig. 1 shows a typical high-power measurement setup. For example, for multipactor tests, the multipactor threshold is the minimum level of incident power ($P_{i,\min}$) on the DUT at which the breakdown takes place. However, this is not the only parameter that needs to be monitored. IL and RL of the DUT must also be known, in order to check for the correct performance of the unit during the test. To do so, both the transmitted and the reflected power need to be properly measured.

Several sources of potential uncertainty and/or errors affecting power measurements have been identified in this article, as shown below. In cases where systematic errors cannot be accounted for and corrected, their effect is to be included in the uncertainty budget.

1) *Single Power Measurement (SPM):* Performed by means of two instruments: power sensor and power meter. The following issues must be assessed.

   a) Measurement instrumentation uncertainty.
   b) Low power versus high power.
   c) Pulsed versus continuous power.
   d) Absolute versus relative measurement.

2) *Test Power Measurement:* Any power measurement stated in a result report of the combination of several SPMs. Hence, the overall uncertainty is a combination of individual contributions (law of propagation of uncertainty).

3) *Adapters:* They are used during the calibration but removed on testing. Errors can arise as a result of their IL, and its effect can be accounted for.

4) *Transient versus stationary regime:* The application of different power levels implies the existence of transient regimes. Conversely, when the nominal testing power is applied and stabilized, a stationary regime is achieved.

5) *Variation of Conductivity:* The setup calibration is carried out at ambient conditions, while the test is performed along a wide range of temperatures, hence affecting conductivity of metals.

6) *Harmonics:* The presence of harmonics along the setup can cause significant error in power readings.

7) *Standing Wave Ratio:* The presence of reflected power in the setup [due to nonideal reflection coefficients (RCs) at different test ports] can affect the test results. Mismatches can also be accounted for in the uncertainty budget.

8) *Frequency Shift:* The setup calibration is carried out at ambient temperature and normal air pressure, while the test is performed in vacuum and at extreme temperatures, hence affecting permittivity and geometry. These physical effects affect the "equivalent" working frequency, as if a frequency shift had taken place.

9) *Change of Physical Position:* The setup calibration is done with the transmission lines deployed in a particular position, but then, when the DUT is connected, this position may remarkably change, affecting the IL.

III. Single Power Measurement

The power sensor and the power meter always work together as a system. When connected to an RF power source, they are used to perform a power measurement. This is what we call it an SPM. Its uncertainty budget can be obtained from the manufacturer operation manuals. Both devices, power sensor and power meter, are usually subjected to a strict periodic calibration schedule to detect any degradation leading to a potential nonoperating status or malfunction. A summary of uncertainty contributions affecting an SPM is given in Table II. The first two columns are as in [34, Tables IV–II]. The third column, added now, shows the assumed probability density function (pdf) for each individual contribution, as well as the divisor to be considered for each pdf. The combined standard deviation is the root-sum-of-squares of all individual contributions, each of them affected by the indicated divisor, which is dependent on the type of pdf considered. By doing this, we are combining all contributions in terms of individual standard deviations. Finally, the expanded uncertainty is obtained by multiplying the combined standard deviation times a coverage factor (usually $k = 2$ for a confidence level of 95.45%).

Every parameter in the list is related to the power sensor and/or the power meter. The exception is M_u, which is dependent also on the RF generator producing the signal being measured and on the rest of instrumentation between the generator and the test port. Particular attention must be paid to P_l, which refers to the linearity of the power sensor.

Linearity is directly related to the low-power versus high-power issue. Generally, the calibration of a power setup is done

TABLE II
Standard Uncertainties for Single Power Measurement [34]

Standard uncertainty	Source	PDF, divisor
u(M_u)	Mismatch gain uncertainty between the sensor and the generator. The standard uncertainty is dependent upon the reflection coefficients of the sensor and the generator. Refer to the mismatch model. Reflection coefficients may have different distributions	U-shaped, sqrt(2)
u(M_{uc})	Mismatch gain uncertainty between the sensor and the calibrator output of the power meter. The standard uncertainty is dependent upon the reflection coefficients of the sensor and the calibrator output. Refer to the mismatch model	U-shaped, sqrt(2)
u(P_m)	Power meter instrumentation uncertainty	Rectangular, sqrt(3)
u(P_{mc})	Power meter instrumentation uncertainty (during calibration)	Rectangular, sqrt(3)
u(D)	Power meter drift uncertainty	Rectangular, sqrt(3)
u(K_b)	Sensor calibration factor uncertainty. Typically, the value of the uncertainty is reported along with the calibration factor by the calibration laboratory or the manufacturer	Gaussian, 2
u(K_c)	Sensor calibration factor uncertainty at the frequency of the power meter calibrator output. If the sensor is calibrated relative to the associated calibrator output frequency, $K_c = 1$ and u(K_c) = 0	Gaussian, 2
u(P_l)	Power sensor linearity, which is related to power range. Generally negligible on lower ranges but has higher uncertainty at high power levels	Rectangular, sqrt(3)
u(P_{cal})	Calibrator output power level uncertainty	Rectangular, sqrt(3)
u(Z_s)	Power meter zero set uncertainty	Rectangular, sqrt(3)
u(Z_c)	Power meter zero carryover uncertainty	Rectangular, sqrt(3)
u(N)	Power meter and sensor noise uncertainty	Rectangular, sqrt(3)

using low-level signals, namely 0–10 dBm. However, during the test, the signal can easily increase up to 50–60 dBm. This implies that the signal level present at the coupler ports ranges from very low levels when calibrating (−30 to −40 dBm) to high levels during the test itself (10–20 dBm). This is a drawback because the measurement uncertainty of power sensors varies as a function of power level: the wider the dynamic range, the higher the uncertainty. This variation of uncertainty as a function of input power is characterized by means of the so-called power sensor linearity. Regarding linearity, two kinds of measurements are specified in the technical literature and specification datasheets: 1) absolute measurement, where
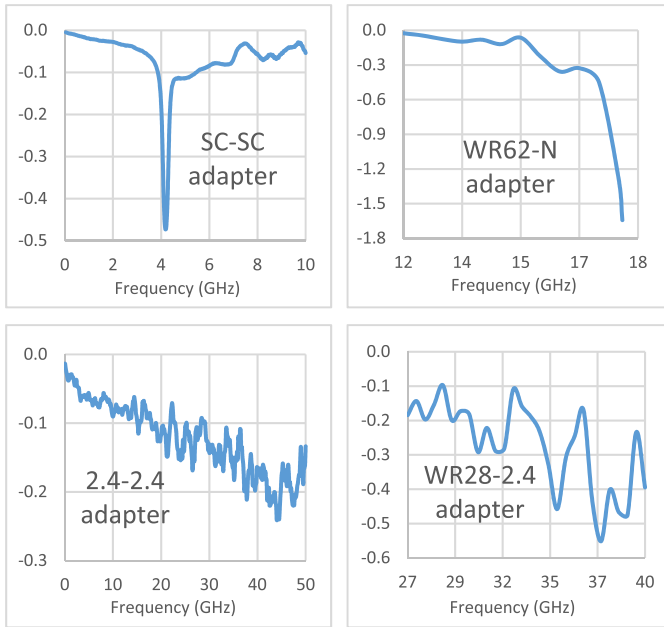
Fig. 2. Insertion loss (IL) in dB (vertical axis) versus frequency for various commercial adapters. Spikes and significant values for IL can be observed, which must be considered in the calibration process for the tests.

a single direct measurement is performed, using only one channel of the power meter; or 2) relative measurement, using both channels at the same time, defined as the relative value (or ratio) between two readings in the power sensors. In relative measurements, linearity can affect remarkably the measurement result, since power levels at both sensors can differ a lot.

In a similar way to the previous case, it is very likely that, during the implementation of a high-power test, both continuous and pulsed power modes need to be used. The discussion on how both modes uncertainty must be dealt with shows up. Effectively, the calibration of a high-power setup is normally done using continuous wave (CW) signals, because the dynamic range of power sensors is wider in the average mode than in the pulsed one. But, on the other hand, there are tests that need to be performed using pulsed power, very common for multipactor and corona tests. Those power sensors capable to measure both types of signals have two different linearity specifications: one for CW signals, another for pulsed ones. However, the difference is negligible, and the uncertainty of both CW and pulsed methods are similar.

An example assigning values to all sources of uncertainty described on Table II was provided by the manufacturer Agilent Technologies [34]. The overall figure for the combined uncertainty (U_c) is 2.43%, i.e., relative to the measured power or power ratio. It is more usual to use the so-called expanded uncertainty (coverage factor $k = 2$ assuming a Gaussian or normal pdf), which gives $U_{exp} = 4.86\%$. It is not specified whether the example of Agilent is absolute or relative or whether the assigned values are either typical or worst case values.

In this article, we have gathered actual values for uncertainty contributions. Data have been extracted from calibration certificates and datasheets for specific power sensors, models

E8487A [35], E9326A [36], and E4413A [37]; for power meters, model E4417A [38]; and for signal generator, model E8257C [39]. The following values for the expanded uncertainty (coverage factor $k = 2$) have been obtained.
1) *Absolute Measurement:* typical 2.6%, worst case 7%.
2) *Relative Measurement:* typical 4%, worst case 7%.

Worst case values are linked to extreme frequencies and/or unlikely instrument states. No relevant difference between CW and pulsed measurements has been found. The relative measurement uncertainty for a typical case is close to Agilent's example (4% versus 4.86%).

IV. TEST POWER MEASUREMENT

A power measurement is actually the combination of several measurements, for example, the incident power (P_i) calibration requires a relative measurement between the interface to which the DUT is connected and the power measuring port (see Fig. 1). Subsequently, during the test, an absolute measurement of power is done at the measuring port. Therefore, the overall uncertainty U_{Pi} associated with any power measurement P_i done in a high-power test (as stated on a test report) is the combination of two uncertainty contributions: 1) the uncertainty associated with the relative measurement (U_{rel}) done during the calibration and 2) the uncertainty associated with the absolute measurement (U_{abs}) done during the test itself.

Note that the usual consensus is to derive and combine all uncertainties in relative magnitude, i.e., as a ratio to the measured value (either an absolute power or a ratio itself). According to the law of propagation of uncertainties [40], [41], the combined uncertainty is the root-sum-of-squares of the two individual contributions, since the sensitivity coefficients are always unity:

$$U_{Pi} = \left(U_{rel}^2 + U_{abs}^2\right)^{1/2}. \tag{1}$$

In the case of reflected power and transmitted power (P_r and P_t, respectively), the analysis is slightly different, since three absolute measurements are necessary.
1) Measurement of a power reference.
2) Measurement of power at DUT interface, necessary to compute the attenuation between the interface the DUT is connected to and the power measuring port.
3) Measurement of power at the measuring port during the test. In this case, the overall uncertainty (U_{Pr} or U_{Pt}) associated with any reflected or transmitted power measurement during a high-power test is the combination of three uncertainty contributions related to three absolute measurements

$$U_{Pr} = U_{Pt} = \left(3 \cdot U_{abs}^2\right)^{1/2}. \tag{2}$$

Typical and worst case values for U_{rel} and U_{abs} were obtained in prior sections. If we introduce their values into (1) and (2), it comes out that during a high-power test, all three power components (P_i, P_r, and P_t) have a very similar uncertainty budget. Their associated expanded uncertainties (coverage factor $k = 2$) can be rounded to the following values.

1) P_i *Power Measurement:* typical 4.8%, worst case 9.9%.
2) P_r *and* P_t *Power Measurement:* typical 4.5%, worst case 12.1%.

Please note that it is normally assumed that individual contributions to the overall uncertainty, which are evaluated at different measurement planes (as it is the case for U_{Pi}, U_{Pr}, and U_{Pt}), are uncorrelated. The reason is that such uncertainties come from the numerous signal reflections produced at each plane, involving the interaction of complex (magnitude and phase) RCs looking at both sides of the considered plane. Since the relative phase between both RCs is usually unknown, the usual approach is to estimate a bound for the mismatch uncertainty, which follows a U-shaped pdf (for this estimation, only the magnitudes of the involved RCs are needed). Under these conditions, it is difficult to foresee any kind of correlation between individual contributions, since they are dependent on the RCs at the diverse measurement planes, whose relative phase is unknown (or considered random for the purpose of the magnitude estimation for the U-shaped pdf). In general, all correlations of this type between individual contributions to the uncertainty have been assumed negligible throughout this article.

It is very common to characterize the performance of a DUT in terms of IL and RL and have the following expressions:

$$IL(dB) = 10 \times \log\left(\frac{P_t}{P_i}\right) \tag{3}$$

$$RL(dB) = 10 \times \log\left(\frac{P_r}{P_i}\right). \tag{4}$$

Since IL and RL are obtained as the combination of two independent measurements, which are considered to be uncorrelated, their associated uncertainties U_{IL} and U_{RL} are given by the following expressions:

$$U_{IL} = \left(U_{Pt}^2 + U_{Pi}^2\right)^{1/2} \tag{5}$$

$$U_{RL} = \left(U_{Pr}^2 + U_{Pi}^2\right)^{1/2}. \tag{6}$$

Introducing here the values obtained for P_i, P_r, and P_t, the following expanded uncertainties (coverage factor $k = 2$) are obtained: IL and RL measurement: typical 6.6%, worst case 15.7%.

It has to be emphasized that the uncertainty contributions considered in this section are unavoidable as they are exclusively due to measurement equipment.

V. ADAPTERS

This kind of passive components transfer signals from one type of connector to another type. To carry out the calibration process, it is necessary to use different kinds of RF adapters, which will afterward be removed during actual testing. In principle, adapters' losses tend to be considered negligible. However, depending on the frequency and the type of connector, they can have considerable IL and even exhibit spiky profiles against frequency. Measurement of some commercial adapters has been done, as shown in Fig. 2 for some typical cases. These adapters serve as examples of the types that are used during the calibration process. Many more can be used depending on the frequency range, on the

kind of necessary interfaces and/or the setup itself. They are a representative example that, effectively, their electrical characterization must be done in order to avoid misassigning losses to the DUT, or any other spurious effects, that are actually coming from the adapters.

Fig. 2 shows that the IL of adapters must be taken into account, since they may potentially lead to errors beyond 25% (1 dB). The effect of adapters used during the calibration process is compensated to reduce their impact on the uncertainty budget to avoid introducing any systematic error. This was achieved by using adapter correction techniques, which involve measuring the effect of adapters on the measurement system and applying a correction to eliminate or reduce their impact on the measurements. In this way, using the most precise characterization equipment (for example, vector network analyzers), the accuracy and reliability of RF measurements can be improved. If losses of adapters are not considered, the most likely errors on measurements are the following: 1) the reading of incident power on DUT is lower than the actual one and 2) the reading of reflected and transmitted power is higher than the actual one. For example, if a waveguide-to-coaxial adapter with IL = 0.5 dB is not accounted for in calculations, in this case, this systematic error propagates onto the characterization of the DUT's IL and multipactor performance.

1) DUT IL can be incorrectly measured 1 dB above its actual value.
2) DUT RL can be incorrectly measured 1 dB below its actual value.
3) DUT multipactor and/or corona threshold can be incorrectly computed 0.5 dB below its actual value.
4) Power delivered to DUT in power handling and PIM tests can be incorrectly computed 0.5 dB below its actual value.

Once the IL of the adapters has been measured and accounted for, it can be considered that this systematic source of error has been corrected for and eliminated. However, measurement of IL of adapter still has an associated uncertainty (U_{ada}), since it has been obtained as the combination of two absolute measurements. This additional uncertainty contribution must be combined with U_{Pi}, U_{Pr}, and U_{Pt} (in general U_{Px}), in order to obtain the final uncertainty budget for a test power measurement (including adapter effect) during the calibration (U'_{Px}):

$$U_{ada} = \left(2 \cdot U_{abs}^2\right)^{1/2} \tag{7}$$

$$U'_{Px} = \left(U_{Px}^2 + U_{ada}^2\right)^{1/2}. \tag{8}$$

Thus, finally, if one adapter is used, the values obtained for the expanded uncertainties (including adapter effect and with coverage factor $k = 2$) can be rounded to the following values.

1) P'_i *Power Measurement:* typical 6.0%, worst case 14.0%.
2) P'_r *and* P'_t *Measurements:* typical 5.8%, worst case 15.7%.
3) *IL' Data:* typical 7.6%, worst case 18.6%.
4) *RL' Data:* typical 7.5%, worst case 18.5%.

These uncertainty budgets would increase if more than one adapter were used; and, remarkably, they are almost impossible
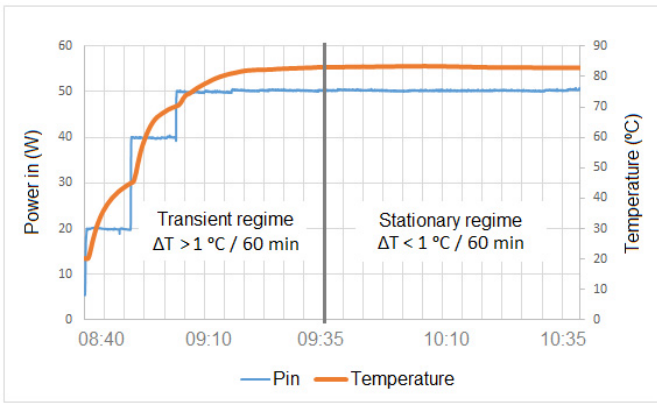
Fig. 3. Two stages of the high-power test: transient regime and stationary regime as a function of time. Temperature variation is due to the application of various power steps. Transient regime between 8:40 and 9:35. Stationary regime between 9:35 and 10:35.

to avoid, since they arise from the inherent uncertainty of the instruments and the unavoidable usage of at least one adapter during the calibration process.

VI. TRANSIENT VERSUS STATIONARY REGIME

The uncertainty assessed in the previous sections is the so-called Type-B uncertainty contribution (U_B). It is based on theoretical or aprioristic knowledge of the measurement setup and/or DUT, extracted from calibration certificates, manufacturer's specifications, previous experience, etc. However, a complete uncertainty assessment must also consider the so-called Type-A uncertainty contributions (U_A), which are based upon the empirical knowledge of the actual measurement, e.g., through repetition of readings to assess for repeatability of measurements or through experimental characterization of specific parameters otherwise obtained from specifications, such as linearity or drift.

Usually, only repeatability is considered as Type-A contribution. Thus, it is noted that this concept is introduced here for the first time. Repeatability of a test is challenging because repeating tests under exactly the same measurement conditions is difficult and time-consuming. However, there is another context in which Type-A evaluation of uncertainty can be of help. It is the assessment of the drift during the stationary regime. A typical high-power test is divided into two successive stages: 1) initial transient regime and 2) final stationary regime. In order to discriminate whether measurements are taken within the first or the second regimes, we normally look at the recorded temperature variation, as in the example of Fig. 3. The most useful outcome of a test is the data extracted from the stationary regime, while data coming from the transient regime is only useful in terms of trend and diagnostic. According to this, this section will only assess data coming from the stationary stage.

In order to estimate the drift, we will apply a least-squares fitting to the data gathered along the stationary regime. Type-A uncertainty contribution (U_A) can be calculated according to the following expression:

$$U_A = \sqrt{(b \cdot T)^2 + 4 \cdot \frac{\sum_{i=1}^{N}(y_i - \langle y_i \rangle)^2}{N - 2}} \qquad (9)$$

TABLE III
ESTIMATION OF DRIFT IN THE STATIONARY REGIME

Avg Value (W)	b·T (W)	σ_y (W)	N	U_A abs (W)	U_A rel (%)
50.272	-0.010	0.223	3,600	0.224	0.44

where b is the slope of the least-squares fitting, N is the number of measurements taken during the stationary regime, and T is its time duration.

The first term inside the square root of (9) is related to the variation in the slope of the least-squares line across the duration of the stationary regime. In an ideal case, it should be zero, since the least-squares fitting of a set of acquired points shows no slope once the stationary conditions are reached.

The second term inside the square root of (9) is the standard deviation of the measured values y_i (the ordinate y is the dependent variable, i.e., power; the abscissa x is the independent value, i.e., time). This second term is an estimation of the dispersion of measured values along the least-squares fit. The ratio is also known as residual variance or unexplained variance: $\langle y_i \rangle$ stand for the adjusted values on the fitting curve, i.e., $\langle y_i \rangle = a + b \cdot x_i$. It is also related to the correlation coefficient R^2 of the least-squares fitting, and it can be considered also a figure of merit of the stationary regime: a set of acquired points ideally distributed along a straight line ($R^2 = 1$) should be the ideal case under stationary conditions.

It is also noted that in other metrological contexts, (9) is used to estimate the time drift of a measurement standard (T being the time until the next re-calibration of the standard), in order to introduce this knowledge into the overall uncertainty budget.

To account for U_A in the final uncertainty budget, U_A must be included as an additional uncertainty contribution in the list stated in Table II, likewise in relative terms. To do that, U_A must be normalized in percentage:

$$U_A(\text{rel}) = \frac{U_A}{\text{Avg Value}} \cdot 100 \qquad (10)$$

where Avg Value is the average value of the power being measured along the stationary regime. An example of Type-A evaluation can be applied to the data shown in Fig. 3. That figure shows a typical high-power test stationary regime of 1-h duration from 09:34 to 10:34. During this period of time, one measurement per second is registered in a table. That means 3600 measurements in an hour: we will take this period of time as T in (9). Fig. 4 and Table III show the change in the slope of the least-squares fitting along one hour (i.e., $b \cdot T$), the standard deviation of the measured values as per the second term in (9), and its residual sum of squares (r.s.s.) combination, i.e., the term U_A.

As it can be noticed, the relative value obtained for U_A is at least one order of magnitude below the rest of Type-B uncertainty contributions assessed in the previous sections. Therefore, this uncertainty contribution can be considered negligible in most cases.
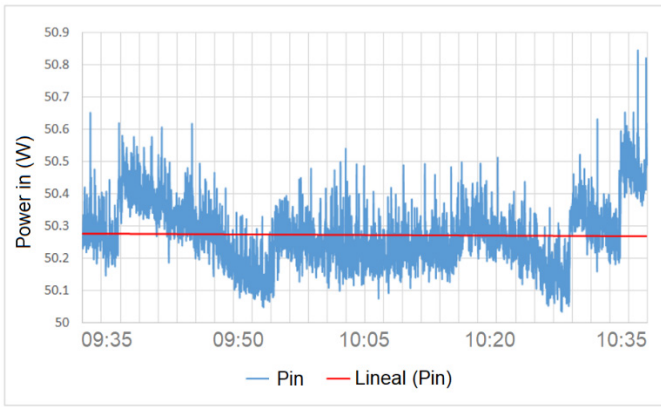
Fig. 4. Least-squares fitting for determination of drift. Stationary regime in the time interval from 09:35 to 10:35.

VII. VARIATION OF CONDUCTIVITY

The setup calibration is performed under ambient conditions, while the actual test is carried out under varying environmental conditions. For example, transmission lines heat up because of Joule losses caused by the limited electrical conductivity (σ) of the materials they are made of [42]. Due to this effect, part of the energy of the high-power signal is lost in the conductive materials of the line. However, heating is not solely caused by the Joule effect but also by the environmental conditions required for the test: it is usual that temperatures over 100 °C need to be applied to the system. On the contrary, sometimes the temperature requirement is very low, such as −100 °C.

ILs of the measurement setup depend directly on conductivity: the higher the conductivity, the lower the ILs of the system (in absolute value). However, conductivity also depends on temperature. This relationship is more conveniently expressed using the inverse of the conductivity, i.e., resistivity (ρ). Its variation depends on both the resistivity of the material and the thermal coefficient of resistance (TCR) and on the temperature variation [43]

$$\rho = \rho_o[1 + \text{TCR} \cdot (T - T_o)]. \qquad (11)$$

Therefore, as temperature increases, both resistivity and ILs increase. Some verifications have been made in order to quantify to which extent the measured ILs of the setup vary during real high-power tests, in relation to the observed variations in temperature. One of these verifications is depicted in Fig. 5, which shows the evolution of the setup's P_i, P_r, and P_t, as well as IL and RL during a high-power test. The test has been carried out on a measurement setup with two 1-m TNC cables inside a thermal vacuum chamber (TVAC), under high vacuum conditions and at the working frequency of 8.5 GHz, with cables' temperature ranging from 20 °C to 120 °C.

The absolute value of IL increased up to 0.5 dB due to heating. Since calibration refers to the point the DUT will be connected to, which is exactly in the middle of the symmetrical line made up of the mentioned two 1-m TNC cables, IL can be assumed to be equally distributed between the input and output paths. This means that the actual incident power on the DUT is 0.25 dB lower than the obtained reading and that the actual transmitted power by the DUT is 0.25 higher than
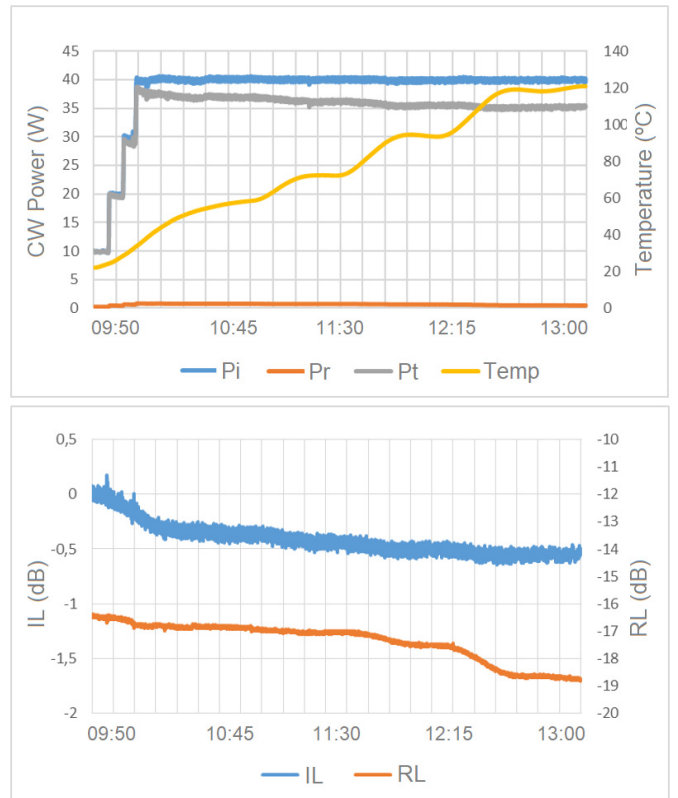

Fig. 5. Evolution of power parameters as a function of time during a test with temperature variation from 20 °C to 120 °C. IL increases by 0.5 dB. Return loss (RL) increases by 2 dB.

the reading. This also means that the actual reflected power by the DUT is 0.25 dB higher than the reading.

The propagation of this systematic error in the characterization of the DUT, in case it is not properly considered, is the following.

1) DUT IL is incorrectly measured 0.5 dB above its actual value.
2) DUT RL is incorrectly measured 0.5 dB below its actual value.
3) DUT multipactor and/or corona threshold is incorrectly computed 0.25 dB above its actual value.
4) Power delivered to DUT in power handling and PIM tests is incorrectly computed 0.25 dB above its actual value.

Other setups were checked under different environmental conditions. Factors such as frequency and technology modify the IL. Summarized results of IL are shown in Table IV. We can see that IL increases with both temperature and frequency.

The mitigation procedure to reduce (or at least account for) this source of uncertainty consists in running a complete power, thermal, and vacuum test on the setup lines, just with the aim of reproducing the final DUT test conditions and therefore to be able to account for the variation of setup performance.

VIII. HARMONICS

High-power amplifiers (HPAs) are a type of equipment prone to produce harmonics. Typical values are around

TABLE IV
INSERTION LOSS VARIATION VERSUS ENVIRONMENTAL CONDITIONS
(RESULTS NOT NORMALIZED)

TEST	Line Type	Freq (GHz)	CW Power (W)	Pressure	Temp variation (°C)	IL variation (dB)
#1	Coax (TNC)	8.5	40	Vacuum	100	0.5
#2	Coax (2.4)	23.0	8	Ambient	No data	0.3
#3	Coax (2.4)	26.0	8	Ambient	No data	0.4
#4	WR42	26.3	90	Vacuum	110	0.3
#5	WR112	7.2	900	Ambient	30	0.1
#6	WR28	37.7	60	Vacuum	50	0.2
#7	WR28	37.7	55	Ambient	No data	0.1
#8	WR28	37.7	55	Vacuum	No data	0.1
#9	WR51	17.0	850	Ambient	No data	0.1

−20 dBc for the second harmonic, but values of −10 dBc and even higher can be easily found [44], [45]. The impact of harmonics on the measurement uncertainty can be very strong, mostly on P_i, since at its measuring point, the harmonic will have experienced little attenuation. For example, if we assume a harmonic incoming of −20 dBc at P_i measuring point: error up to 21% (0.83 dB) can take place

$$P_i = 100 \cdot P_{\text{harm}} \tag{12}$$

$$V_i = 10 \cdot V_{\text{harm}} \tag{13}$$

$$V_i' = V_i \pm V_{\text{harm}} \tag{14}$$

$$0.9V_i \leq V_i' \leq 1.1V_i \tag{15}$$

$$0.81P_i \leq P_i' \leq 1.21P_i. \tag{16}$$

This effect can be enhanced depending on the power sensor technology. If diode technology is used, when working above the square law region, the harmonic effect can cause errors up to 0.9 dB [34].

The mitigation procedure for this kind of uncertainty consists in using low-harmonic amplifiers, besides low-pass filters (LPFs) placed just at the amplifier output (see LPF at Fig. 1). It is also good practice to perform a correlation between incident and transmitted power on the setup, using, in case of incoherence, the transmitted power as a reference, as it is considered that the harmonic will be more attenuated at that point.

IX. STANDING WAVE RATIO

Impedance mismatch is a major issue in any RF high-power test bench. This subject is clearly considered in the ECSS standard [22], where multipactor security margin is linked to payload mismatch. This is consistent with the fact that multipactor is a peak-power-dependent effect and that stationary waves resulting from the reflected signal can produce an important increase in power peaks.

But apart from peak power effects, other points must be assessed regarding reflected power. It must be noticed that

reflected power is not a single component signal. It is the result of the combined interference of all reflected signals along the setup, including those from the DUT. This interference is dynamic, comprised within a changing range of values related to the change in electrical paths according to environmental conditions. Fig. 5 is an example of this: initially a 0.5 dB increase in the absolute value of RL had been predicted due to the increase in IL (as explained in prior sections). However, after a given time, around 12:10 hours, and although the variation of IL has almost stopped, RL exhibits an additional variation of 1.5 dB, in parallel with the last temperature step. This can only be explained as the result of the variating interference of the various reflected components in function of temperature. And even more: if there were a dominant reflected power component coming from the setup, it may mask the power reflected by the DUT itself, making the test outcome hardly useful.

Another drawback coming from the presence of an important amount of reflected power from the setup is the uncertainty that may impact on incident and transmitted power measurements through coupler's directivity. For example, assuming a −10 dBc reflected component and −20 dBc coupler directivity, in a similar case as harmonics, following (12)–(16), a 2% uncertainty can propagate to P_i and P_t readings.

In summary, during the calibration process, IL due to mismatch between setup's elements is accounted for. The issue arises if a reflected component is higher than −15 dBc. The mitigation procedure for this kind of uncertainties is to keep standing wave ratio (VSWR) within safe limits, mostly by choosing the right equipment for the required test specifications. If not possible, like in an out-of-band test, it is good practice to dispose some alternative references for stating power levels.

X. FREQUENCY SHIFT

It has been already said that calibration takes place at ambient conditions, whereas subsequent tests are performed in vacuum conditions at variable temperatures. Due to these changes, permittivity and geometry of the devices [46], [47] involved in the test can also vary, leading to a change in the RF response.

When moving from air during the calibration to vacuum for the test, permittivity slightly changes. Free air relative permittivity (ε_r) is 1.0006, whereas in vacuum its value is exactly 1. This means that the wavelength associated with any RF signal propagating in air is 0.03% shorter than in vacuum (18). It is as if frequency had shifted 0.03% upward. For example, for a 10 GHz signal, this frequency shift is 30 MHz to higher values. We denote this effect as frequency shift

$$v = \lambda \cdot f = \frac{1}{\sqrt{\varepsilon_r \varepsilon_o \mu_r \mu_o}} \tag{17}$$

$$\frac{\lambda_{\text{air}}}{\lambda_{\text{vac}}} = \frac{1}{\sqrt{\varepsilon_r}} = 0.9997. \tag{18}$$

A second factor is the geometry variation of the transmission lines of the setup due to temperature changes. Transmission

lines heat up because of Joule losses as well as from the environmental requirements for the test, as explained in prior sections. The system expands or contracts depending on temperature. When heated up, every dimension is increased: for example, the height and width of a waveguide [46], [47]. For example, the variation of length is dependent both on the material's thermal expansion coefficient (TEC) and on the temperature variation (ΔT) [43]

$$\Delta l = \text{TEC} \cdot l \cdot \Delta T. \tag{19}$$

In this case, for example, a WR112 waveguide, whose operative bandwidth runs from 10 to 15 GHz, with an approximate window size of 30×15 mm, made of aluminum (TEC $= 24 \times 10^{-6}$ $°C^{-1}$), under a temperature variation of $+100$ °C, will experience a geometry expansion up to 0.2% with respect to ambient conditions. It is as if the wavelength of the RF signal had decreased by 0.2% with respect to the dimensions of the transmission system, or just as if the frequency had shifted upward. Conversely, if the system is cooled down, all dimensions contract, thus resulting in a frequency shifting downward. For the specific case study we are presenting, a frequency shift between 10 and 30 MHz may be encountered.

The significance of this phenomenon for the test output is as follows. Normally, a high-power test setup is made of wideband equipment, whose main characteristic is its performance flatness. This points to a low-profile impact. However, a variation of a few tenths of dB in the measured IL may easily arise as a consequence of frequency shift.

A mitigation procedure to account for frequency shift would consist in measuring the S-parameters of the setup's section that will be under vacuum and extreme temperature conditions, in order to assure that, within a bandwidth of several tenths of MHz around the carrier, there are no remarkable spikes. Another option is to run a complete power, thermal, and vacuum test on the setup, just as mentioned in prior sections, so that any unwanted variations could be accounted for.

The frequency shift phenomenon may also be affected by potential instabilities on the environmental conditions during the testing, namely temperature, pressure, and vibration. The authors have taken these factors into consideration and evaluated their potential impact on the measurements while calculating the uncertainty budget. Regarding temperature, an accurate control of the temperature stability of the thermal chambers used during this article activities, has been carried out. To do so, in our systems, the temperature is controlled, and its variation has been kept below 1 °C. Therefore, the goal of achieving negligible impact on measurement uncertainties was achieved.

In relation to the impact of pressure variation, its instability during the execution of the tests is extremely little, around 10%, compared to the transition from ambient pressure to high vacuum conditions, which is around eight to nine orders of magnitude lower than the atmospheric pressure.

Finally, it is certainly true that vibration can cause mechanical noise and affect the stability of the measurement setup. But, in the framework of this article, it can be stated that TVACs are heavy machinery, firmly attached to the ground, to which the

TABLE V
INSERTION LOSSES FOR VARIOUS BENDING ANGLES

Interface	Line type	Bandwidth measured (GHz)	IL variation (dB/m)		
			90⁰	180⁰	Others worst-case
WR112	Waveguide	7 - 10	< 0.1	< 0.1	< 0.1
WR90	Waveguide	8.2 - 12.4	< 0.1	< 0.1	< 0.1
WR75	Waveguide	10 - 15	< 0.1	< 0.1	0.15
WR51	Waveguide	15 - 22	< 0.1	< 0.1	< 0.1
WR42	Waveguide	18 - 26.5	< 0.1	< 0.1	< 0.1
WR28	Waveguide	26.5 - 40	< 0.1	< 0.1	0.20
7/16	Coaxial	0 - 7.5	< 0.1	0.45	0.50
N - N (#1)	Coaxial	0 - 18	0.90	0.70	2.00
N - N (#2)	Coaxial	0 - 18	0.15	0.10	0.50
N - N (#3)	Coaxial	0 - 18	0.20	0.15	0.20
SMA - SMA	Coaxial	0 - 18	0.25	0.30	0.25
3.5 - 3.5	Coaxial	0 - 26.5	< 0.1	< 0.1	< 0.1
2.9 - 2.9 (#1)	Coaxial	0 - 40	0.55	0.50	0.55
2.9 - 2.9 (#2)	Coaxial	0 - 40	< 0.1	< 0.1	0.20
2.4 - 2.4 (#1)	Coaxial	0 - 50	1.30	0.90	1.70
2.4 - 2.4 (#2)	Coaxial	0 - 50	< 0.1	< 0.1	< 0.1

RF setup is also strongly joined. No other significant effect has been identified during testing activities that might potentially be consequence of physical vibrations.

As a result, the authors have determined that these potential instability sources, affecting temperature, pressure, and vibration, have negligible contribution in comparison with the other considered sources to the final uncertainty budget.

XI. CHANGE OF PHYSICAL POSITION

We consider here the effect of changing the physical position of the transmission line and/or cable bending between the calibration process and the test itself. Again, the calibration is done with a particular position of the transmission lines, but when the DUT is connected, this position can remarkably change (mostly when the DUT is of large dimensions). Do IL vary as a consequence of changes in physical position?

To estimate the variation in ILs as a function of the particular physical position of a transmission line, a set of measurements on flexible waveguides and coaxial lines has been carried out. The transmission line got fixed at six different positions, with known bending angles: 0° (straight position), 45°, 90°, 180°, 270°, and 360°. ILs were measured for each case study. The reference IL obtained for the straight position was then subtracted from other positions' IL in order to identify the variations due to position change. Data were normalized to dB/m. The results are summarized in Table V. Measurement uncertainty corresponds to two combined absolute measurements (see prior sections).

We can see that flexible waveguide lines are much more stable than coaxial ones. For usual bending angles such as

90° or 180°, errors up to 1.30 dB can be encountered in coaxial lines. There are extreme differences between coaxial technologies. Errors are observed to increase with frequency. Despite IL variation could be considered a systematic error and therefore corrected for, it might be more appropriate to consider it an uncertainty contribution, because it is very difficult to state beforehand what particular angle variation the line experiences between calibration and test. And even more, if we recall prior sections, measurement of P_r and P_t required a first absolute measurement to obtain a known power reference. This measurement is usually taken by means of a coaxial cable. Should this cable show amplitude instability, as some of those reflected on Table V, the systematic error/uncertainty contribution affecting power measurement might be significant and might compromise the whole test, either if it is carried out on waveguide lines or in coaxial.

Mitigation actions for this kind of variation consist in trying to keep as much as possible the same line positions both for the calibration and for the test. Moreover, it is a good practice to use waveguide lines as much as possible. Finally, when the use of coaxial lines is unavoidable, the proper technology must be chosen so that they can be amplitude stable.

XII. Conclusion

This article is devoted to obtaining a complete uncertainty budget for RF high-power testing activities and to assess whether it may have impact on test security margins, a subject that had never been covered to this extend so far to the best of the authors' knowledge. A wide range of uncertainty contributions and systematic errors has been assessed. As a result, a complete and traceable uncertainty budget for the purposes of evaluating security margins has been derived.

Some uncertainty sources are unavoidable (such as those coming from the specifications of instruments, which are in the order of 5% uncertainty). Others have been proved to be of negligible impact (such as the high-power versus low-power issue, and the transitory regime versus stable regime). Finally, other uncertainty contributions, despite being remarkable (for example, 1 dB due to the use of adapters), can be mitigated or completely removed by taking specific actions (selection of equipment, previous studies, and/or measurements, etc.).

Because of all the above, the measurement uncertainty and systematic errors may eventually affect the results of the characterization of a DUT. If they are not carefully accounted for, they can have an important impact on test results in terms of multipactor threshold, corona threshold, IL, RL, and passive intermodulation performance. In fact, a statement of associated uncertainty to every parameter under test as shown in this article can be included in the test reports for completeness.

Acknowledgment

References

[1] *Testing*, Standard ECSS-E-ST-10-03C, ECSS Secretariat, Space Engineering, ESA-ESTEC Requirements & Standards Division, 2012.

[2] T. Graves, "Standard/handbook for radio frequency (RF) breakdown prevention in spacecraft components," Space Missile Syst. Center, Air Force Space Command, El Segundo, CA, USA, AEROSPACE Rep. TOR-2014-02198, 2014.

[3] P. Partridge, "Standard/handbook for RF ionization breakdown prevention in spacecraft components," Space Missile Syst. Center, Air Force Space Command, El Segundo, CA, USA, AEROSPACE Rep. TOR-2015-02546, 2015.

[4] A. J. Hatch and H. B. Williams, "The secondary electron resonance mechanism of low-pressure high-frequency gas breakdown," *J. Appl. Phys.*, vol. 25, no. 4, pp. 417–423, Apr. 1954.

[5] J. R. M. Vaughan, "Multipactor," *IEEE Trans. Electron Devices*, vol. 35, no. 7, pp. 1172–1180, Jul. 1988.

[6] J. R. M. Vaughan, "A new formula for secondary emission yield," *IEEE Trans. Electron Devices*, vol. 36, no. 9, pp. 1963–1967, Sep. 1989.

[7] J. Petit, "Diagnostic investigations into the multipactor effect, susceptibility zone measurements and parameters affecting a discharge," ESTEC, Noordwijk, The Netherlands, Working Paper 1556, 1989.

[8] F. Höhn, W. Jacob, R. Beckmann, and R. Wilhelm, "The transition of a multipactor to a low-pressure gas discharge," *Phys. Plasmas*, vol. 4, no. 4, pp. 940–944, Apr. 1997.

[9] R. Woo, "RF voltage breakdown in coaxial transmission lines," *Proc. IEEE*, vol. 57, no. 2, pp. 254–256, Feb. 1969.

[10] R. Woo, "Final report on RF voltage breakdown in coaxial transmission lines," Jet Propuls. Lab., Pasadena, CA, USA, Tech. Rep. 32-1500, 1970.

[11] J. M. Meek, *Electrical Breakdown of Gases*. London, U.K.: Oxford Univ. Press, 1953.

[12] D. Smacchia, "Comparison of third-order passive inter-modulation products in two and three high power RF scenarios," in *Proc. MULCOPIM*, 2017, p. 7.

[13] C. Vicente, "Passive intermodulation and corona discharge for microwave structures in communications satellites," Ph.D. dissertation, Elektrotechnik und Informationstechnik der Technischen Universitat Darmstadt, Darmstadt, Germany, 2005.

[14] V. Kuronen, "Passive intermodulation tester characteristics," M.S. thesis, Fac. Inf. Technol. Elect. Eng., Univ. Oulu, Oulu, Finland, 2020.

[15] *Understanding PIM, Part No 11410-00629F*, ANRITSU Company, Bengaluru, India, 2014.

[16] F. Zimmermann, D. Schulte, R. Cimino, C. Vaccarezza, M. Zobov, and R. Wanzenberg, "Electron cloud in wigglers," Conseil Européen pour la Recherche Nucléaire (CERN), Meyrin, Switzerland, Tech. Rep. CERN-OPEN-2006-005; CLIC-Note-650; EUROTeV-Rep.-2006-002, 2006.

[17] I. Montero et al., "Low-secondary electron emission yield under electron bombardment of microstructured surfaces, looking for multipactor effect suppression," *J. Electron Spectrosc. Rel. Phenomena*, vol. 241, May 2020, Art. no. 146822.

[18] A. Ruiz et al., "UHV reactive evaporation growth of titanium nitride thin films, looking for multipactor effect suppression in space applications," *Vacuum*, vol. 81, nos. 11–12, pp. 1493–1497, Aug. 2007.

[19] M. Mostajeran, E. T. Tulu, and U. van Rienen, "Uncertainty in the isosceles multipactor threshold of triangularly grooved surfaces based on polynomial chaos," *Nucl. Instrum. Methods Phys. Res. A, Accel. Spectrom. Detect. Assoc. Equip.*, vol. 993, Mar. 2021, Art. no. 165001.

[20] S. Anza et al., "Novel prediction methods of multicarrier multipactor for space industry standards," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 1, pp. 670–684, Jan. 2022.

[21] R. Udiljak, D. Anderson, M. Lisak, V. E. Semenov, and J. Puech, "Multipactor in a coaxial transmission line. I. Analytical study," *Phys. Plasmas*, vol. 14, no. 3, Mar. 2007, Art. no. 033508.

[22] *Multipactor Design and Test*, Standard ECSS-E-ST-20-01C, Space Engineering, ECSS Secretariat, ESA-ESTEC Requirements & Standards Division, 2020.

[23] *Electrical and Electronic*, Standard ECSS-E-ST-20C, Space Engineering, ECSS Secretariat, ESA-ESTEC Requirements & Standards Division, 2008.

[24] *Eurospace Consortium Technology Harmonisation Panel (THP), Power RF Measurement and Modelling, ESA Harmonisation Mapping*, Eur. Space Res. Technol. Centre (ESTEC), Eur. Space Agency (ESA), Noordwijk, The Netherlands, 2015.

[25] D. Raboso, "Power RF measurements & modelling," IPC-THAG Eur. Space Technol. Harmonisation Dossier, Eur. Space Res. Technol. Centre (ESTEC), Eur. Space Agency (ESA), Noordwijk, The Netherlands, Tech. Rep., 2021.

[26] *International Vocabulary of Metrology—Basic and General Concepts and Associated Terms*, document JCGM 200:2012, JCGM, 2012.

[27] Fluke. *Introduction to the ISO Guide to the Expression of Uncertainty of Measurement (GUM)*. Accessed: 2022. [Online]. Available: https://us.flukecal.com/blog/introduction-iso-guide-expression-uncertainty-measurement-gum

[28] *ISO Guide to the Expression of Uncertainty in Measurement*, Int. Org. Standardization, Geneva, Switzerland, 1995.

[29] A. Ferrero et al., "Uncertainty of material measurements at microwave frequencies," in *Proc. IEEE 8th Int. Workshop Metrol. Aerosp. (MetroAeroSpace)*, Naples, Italy, Jun. 2021, pp. 232–234, doi: 10.1109/MetroAeroSpace51421.2021.9511768.

[30] Z. Hao, "Error analysis and uncertainty evaluation in microwave high power measurement," in *Proc. Int. Conf. Microw. Millim. Wave Technol.*, Chengdu, China, May 2010, pp. 898–901, doi: 10.1109/ICMMT.2010.5525160.

[31] *Passive RF and Microwave Devices, Intermodulation Level Measurement—Part 1: General Requirements and Measuring Methods*, Standard IEC 62037-1:2012, International Electrotechnical Commission, CENELEC, 2012.

[32] *Requirements for the Control of Electromagnetic Interference Characteristics of Subsystems and Equipment*, Standard MIL-STD-451G, Department of Defense Interface Standard, 2007.

[33] K. Kuhlman et al., "Metrology for industrial connectors at radio frequencies—METCONRF," EMPIR Potential Res. Topic, Eur. Metrol. Programme Innov. Res. (EMPIR), Eur. Assoc. Nat. Metrol. Inst. (EURAMET), Braunschweig, Germany, Rep. PRT2017 IND ELEC 071, 2017.

[34] H. Packard, "Fundamentals of RF and microwave power measurements," AGILENT Technol., Chennai, India, Appl. Note 64-1A, 1998.

[35] *8480 Series Coaxial Power Sensors Operating and Service Manual, Part Number 08481–90173*, Keysight, Santa Rosa, CA, USA, 2017.

[36] *E-Series E9320 Operating and Service Guide, Part Number E9321–90001*, Keysight, Santa Rosa, CA, USA, 2011.

[37] *E4412A and E4413A Operating and Service Guide, Part Number E4412–90013*, Keysight, Santa Rosa, CA, USA, 2016.

[38] *Agilent E4416A/E4417A EPM-P Series Power Meters and E-Series E9320 Peak and Average Power Sensors Data Sheet, Part Number 5980–1469*, AGILENT Technol., Chennai, India, 2009.

[39] *Agilent E8247C/57C PSG CW and Analog Signal Generators Data Sheet, Part Number 5988–7454*, AGILENT Technol., Chennai, India, 2007.

[40] B. Taylor, "Guidelines for evaluating and expressing the uncertainty of NIST measurement results," NIST, Gaithersburg, MD, USA, Tech. Note 1297, 1994.

[41] *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement*, document JCGM 100:2008, 2008.

[42] R. E. Collin, *Field Theory of Guided Waves*. New York, NY, USA: IEEE Press, 1991.

[43] R. Serway, *Física*. New York, NY, USA: McGraw-Hill, 1992.

[44] Bonn Elektronik. *TWAP Series Amplifiers Specifications*. Accessed: 2022. [Online]. Available: https://www.bonn-elektronik.com/en/RF-power-amplifiers/twap-series/

[45] Bonn Elektronik. *TWAL Series Amplifiers Specifications*. Accessed: 2022. [Online]. Available: https://www.bonn-elektronik.com/en/RF-power-amplifiers/twal-series/

[46] J. Uher, *Waveguide Components for Antenna Feed Systems*. Boston, MA, USA: Artech House, 1993.

[47] R. Cameron, *Microwave Filters: Fundamentals, Design and Applications*. Hoboken, NJ, USA: Wiley, 2007.

Martín García-Patrón received the B.Sc. degree in telecom engineering from the Universidad de Alcalá de Henares (UAH), Madrid, Spain, in 1998 and the M.Sc. degree in applied physics from the Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain, in 2009.

Since 2005, he has been with the Spanish National Institute for Aerospace Technology (INTA), Madrid, Spain, where he was appointed as a Manager of the RF High Power Laboratory in 2019. He is an expert on multipaction and passive intermodulation testing for space-borne equipment. He is involved in different international forums devoted to improving technology and testing procedures (EURAMET FutureCom, MICIN Generación del Conocimiento, ESA MULCOPIM, and IEC TC46).

Manuel Rodríguez received the B.Sc. degree in telecommunications engineering from the Universidad Politécnica de Madrid, Madrid, Spain, in 1989.

He has been with the Metrology and Calibration Centre, Instituto Nacional de Técnica Aeroespacial (INTA), since 1992, and has been in charge of the RF and Microwave Laboratory since 2015. By agreement with the Spanish Centre of Metrology (CEM), it acts as Designated Institute (DI) of the Spanish national standards in the following RF magnitudes: power, impedance, electromagnetic noise, and attenuation. He has also joined several EMRP and EMPIR projects, such as 'Ultrafast' and 'RFMicrowave'.

Jorge A. Ruiz-Cruz (Senior Member, IEEE) received the Ingeniero de Telecomunicación and Ph.D. degrees from the Universidad Politécnica de Madrid, Madrid, Spain, in 1999 and 2005, respectively.

Since 2006, he has been with the Universidad Autónoma de Madrid, Madrid, where he was appointed as an Associate Professor in 2009 and a Full Professor in 2021. His current research interests include the computer-aided design of microwave passive devices and circuits (filters, multiplexers, and orthomodes).

Isabel Montero is currently a Research Professor with the Spanish National Research Council (CSIC), Madrid, Spain. She is also the Head of the Surface Nanostructuring for Space and Terrestrial Communications Group, Materials Science Institute of Madrid (ICMM-CSIC), Madrid, Spain, and the Director of the Spanish Physics Laboratory on Secondary Electron Emission (SEY), CSIC. She is an expert in surface spectroscopic techniques and in low-secondary electron-emission surfaces and coatings to avoid multipaction effect.

Prof. Montero is a Working Group Member of the ECSS Multipactor Working Group ECCS-E-ST-20-01-Multipaction, ASD-EUROSPACE.