Forward-stepwise regression analysis for fine leak batch testing of wafer-level hermetic MEMS packages

Changsoo Jang a, Byeng Dong Youn a, Ping F. Wang a, Bongtae Han a, Suk-Jin Ham b

Abstract
An advanced regression scheme is proposed to analyze fine leak batch testing data of multiple MEMS packages. The scheme employs the forward-stepwise regression method to infer the information of leaky packages from a batch test data. The analysis predicts the number of leaky packages and the true leak rate of each leaky package in a progressive manner. The scheme is implemented successfully using an actual batch test data obtained from wafer-level hermetic MEMS packages. An error analysis is followed to define the applicable domain of the scheme. Advanced formulations are also suggested to extend the applicable domain.

1. Introduction
The helium mass spectrometer based leak test is widely used in the industry for fine leak detection [1,2]. The output of the spectrometer is an apparent leak rate, which is defined as the leak rate of a given package as measured under specified test conditions [3]. In practice only the initial apparent leak rate is used as a qualitative measure of hermeticity.

An individual true leak test is essential to inspection of the hermetic performance of each package. The information of individual leaky packages is often required when a fabrication process is to be evaluated. However, such an individual test is not practical in a mass production environment because of long testing time. A test with multiple packages is a preferred choice in the industry, where a batch of packages is subjected to the same test conditions. Yet, the current procedure of batch tests can provide only qualitative information about possible leaky packages in the batch, which is not sufficient to address the issues associated with manufacturing yields.

In this paper we extend the He fine leak test based true leak rate measurement into the batch test domain. The scheme employs the forward-stepwise regression method to determine the number of leaky packages and their true leak rates from a single batch test data.

2. Background: gas leak rate of an individual package
In the metallic seals of wafer-level hermetic MEMS packages, gas transport occurs through randomly present nano-scale leak channels and thus is governed by the molecular gas conduction theory [4]. Even when multiple nano-scale leak paths exist, they can be modeled as an effective single leak channel. The leak rate depends on the gas molar mass and the geometry of the channel (diameter and length). The conduction-based governing equation for leakage, known as Howl–Mann equation [5], is expressed as

\[
R_i = \frac{L_P}{P_0} \left[ 1 - \exp\left( \frac{-L_V}{VP_0} \right) \right] \exp\left( \frac{-L_{dwell}}{VP_0} \right)
\]

where \(R_i\) is the initial apparent leak rate obtained at the beginning of spectrometer operation (atm-cc/s); \(L\) is the true leak rate of a gas (atm-cc/s); \(L_P\) is the bombing time (s) during which a package is subjected to helium pressurized at the bombing pressure; \(P_0\); \(L_{dwell}\) is the dwell time (s) which is the short duration between the instant the specimen is taken out of the bombing chamber and the instant the spectrometer is switched on; \(V\) is the package cavity volume (cc); and \(P_0\) is the standard pressure (=1 atm). The existing standard [3] utilizes the equation for hermeticity qualification. It should be noted that the square root ratio between molar masses of helium and air, which is included in the original Howl–Mann equation to convert a helium leak rate to a true air leak rate (\(L_a\)), is omitted in Eq. (1) because only the helium leak rate is considered in this study.

As can be seen from Eq. (1), the initial apparent leak rate can vary significantly with testing parameters. Therefore, the standard using the initial apparent leak rate can serve only as a qualitative benchmark for the small cavity volumes.
Using the original Howl–Mann equation, the apparent leak rate, \( R \) (atm-cc/s), at the end of the zero signal time, \( t_{zero} \), can be expressed as

\[
\Omega = \frac{L_{Pb}}{P_0} \cdot t \cdot \exp \left( \frac{-L_{tb}}{VP_0} \right) \cdot \exp \left( -\frac{L_{t_{dwell}} + t_{zero}}{VP_0} \right)
\]

where \( R \) is the apparent leak rate (atm-cc/s) and \( t \) is the elapsed time during the measurement phase (s). The above equation describes well the apparent leak rate history since the gas flow inside a spectrometer chamber is always in the molecular conduction regime [4]. Then the true leak rate can be calculated accurately through a regression analysis [6].

Another important practical aspect of the fine leak test is the measurement range associated with the spectrometer sensitivity. Due to the finite helium spectrometer sensitivity, the fine leak test cannot detect true leak rates below or above certain limits. Eq. (2) is used to illustrate the limits. In Fig. 2, the initial leak rate is plotted as the apparent leak rate during the measurement phase can be simplified as

\[
R(t) = \Omega \exp \left( -\frac{Lt}{VP_0} \right)
\]

where \( R \) is the apparent leak rate (atm-cc/s) and \( t \) is the elapsed time during the measurement phase (s). The above equation describes well the apparent leak rate history since the gas flow inside a spectrometer chamber is always in the molecular conduction regime [4]. Then the true leak rate can be calculated accurately through a regression analysis [6].

In Fig. 2, the initial leak rate is plotted as the function of true leak rates. The parameters used in the calculation were: \( P_0 = 5 \) atm, \( t_0 = 6 \) h, \( V = 2.156 \times 10^{-6} \) cc, \( t_{dwell} = 10 \) min and \( t_{zero} = 150 \) s. The plot shows that the detectable true leak rate ranges from \( 4.5 \times 10^{-10} \) atm-cc/s to \( 3.5 \times 10^{-8} \) atm-cc/s if the spectrometer sensitivity is \( 10^{-10} \) atm-cc/s. In practice this range is reduced further due to measurement uncertainties associated with the data acquisition system.

3. Methodology

The idea of inferring the information about leaky packages from a batch test result is based on the fact that each leaky package generates a unique signal pattern defined by Eq. (3), and accordingly, the total signal should also have a unique pattern. A total apparent leak rate of \( n \) leaky packages \( (R_0) \) can be expressed as

\[
R_0(t) = \sum_{i=1}^{n} \Omega_i \exp \left( -\frac{L_i t}{VP_0} \right)
\]

The unknowns in the above equation are the number of leaky packages \( n \), and the initial leak rate \( (\Omega_i) \) and true leak rate \( (L_i) \) of each leaky package. This total signal can be decomposed into the original set of individual signals using a proper data processing technique. The forward-stepwise regression method is employed to conduct this multi-parameter, non-linear inverse analysis.

A forward-stepwise regression analysis [7] uses an additive form of exponential functions [Eq. (4)] to model the apparent leak rate data. The analysis predicts the number of leaky packages \( n \), the initial leak rate \( (\Omega_i, i = 1 \text{ to } n) \), and the true leak rate \( (L_i, i = 1 \text{ to } n) \) in a progressive manner. The regression analysis begins with \( n = 1 \) and increases \( n \) gradually until a convergence criterion is met.

3.1. Regression criterion

The prediction square sum error is used as a convergence criterion to ensure that the prediction accuracy is achieved to a desired level. The normalized \( L^2 \) regression error norm \( (\varepsilon) \) is defined mathematically as

\[
\varepsilon_t = \frac{1}{\sqrt{n}} \sum_{t=0}^{t_{max}} \left( \frac{R(t) - \hat{R}(t)}{\mu_{\Omega}} \right)^2 \quad \text{and} \quad \mu_{\Omega} = \frac{1}{n} \sum_{t=0}^{t_{max}} R(t)
\]

where \( \hat{R}(t) \) is the regression model of the apparent leak rate, \( n \) is the number of data points \( (= t_f/\Delta t + 1; \Delta t \) and \( \Delta \) are the time duration and interval of data acquisition, respectively) and \( \mu_{\Omega} \) is the mean of the observed apparent leak rate.

In general, a smaller regression error norm would provide higher convergence accuracy. If it is too small, however, the analysis may not converge. The error norm should be determined carefully considering how well the mathematical function describes the physical phenomenon (i.e., the assumptions used in the mathematical model) and the level of noise in experimental measurements. More importantly, the non-linear regression analysis with multiple roots can produce values which are not physically valid. Extra criteria are required to reject invalid solutions.
3.2. Extra criteria for physical

The following extra criteria are proposed to avoid mathematically feasible but physically invalid roots and thus ensure the convergence of the regression analysis in a strong sense. The extra criteria are developed from three physical conditions that should not be violated: (1) the true and initial leak rates are positive, (2) the resolution of the helium spectrometer is limited and the apparent leak rate should be larger than the spectrometer sensitivity limit, and (3) there should be one-to-one correspondence between the true leak rate and the apparent leak rate. The first criterion is simple to implement and the criteria corresponding to the last two conditions are described below.

(1) Minimum value for the initial leak rate ($e_{II}$):

$$ e_{II} = \min(\Omega_0) > e_{II} $$

This regression analysis takes the minimum $\Omega_0$ value as a convergence indicator for the second physical condition. When $\Omega_0 \leq e_{II}$, a package is regarded as hermetic. The value of $e_{II}$ should be the minimum initial leak rate that can be captured by the helium spectrometer.

(2) Minimum difference of the true leak rate values ($e_{III}$):

$$ e_{III} = \min_j |\hat{L}_j(t) - \hat{L}_j(t)| > e_{III}, \quad \forall i, j \in [0, n] $$

The difference between any of two predicted values for a true leak rate can be used as another convergence indication for the third physical condition. Numerical noise in the apparent leak rate may lead to an artificial increase in the number of leaky packages ($n$) by numerically splitting one leaky package into multiple artificial ones with the same or numerically equivalent leak rate but different $\Omega$ values. The criterion $e_{III}$ should correspond to the minimum measurable true leak rate.

4. Implementation

The proposed method was implemented for an actual batch test data. The package and the experimental procedure are described briefly. The batch test results are presented first and the validation test results are followed.

4.1. Package and test procedure

The wafer-level hermetic MEMS package used in the test is shown schematically in Fig. 4. The seal was made by the Sn–Au eutectic bonding process and the width of the metallic seal was 0.13 mm.

A flow chart of the proposed regression analysis is shown in Fig. 3. The regression process ends when the regression criterion [Eq. (5)] is first met. If the second and third criteria are also met, the analysis is completed. If either of the extra criteria is violated, the initial values of the regression are changed and a new regression analysis is performed with the known number of package. This trial step is repeated until all the criteria are met or the number of trials reaches a certain number ($j_e$ in Fig. 3). If the results do not satisfy all the criteria at the end of $j_e$ trials, the analysis ends and the results at the previous step ($n - 1$ packages) becomes the final solution.

![Fig. 3. Flow chart of the proposed regression analysis procedure.](image-url)

![Fig. 4. Schematic of test MEMS package.](image-url)
We followed a standard leak test procedure described in Fig. 1, which was implemented for an individual package in Ref. [4]. In each leak test, the signals (apparent leak rates) were documented continuously for 1–2 h after 6 h of bombing at 5 atm and 10 min of dwell time for transferring specimens from the bombing chamber to the spectrometer chamber. Details about the fine leak test procedure and experimental setup can be found in [8].

Fig. 5 illustrates the entire test and regression procedure that was performed in this study. A total of 100 packages were fabricated and it was confirmed through the first-level batch test (Batch 1) that there should be leak package(s). The signal obtained from Batch 1 was used for a regression analysis and the information about leaky packages was attained. In order to verify the regression result, extra batch tests as well as individual tests were performed. In the second-level test the original batch was split into two sub-batches (Batch 1-1 and 1-2). The sub-batches were subjected to the same leak testing and only those with evident leakage were selected for the next level batch tests. Such hierarchical testing continued until we obtained complete leakage information of each leaky package. As a result, 28 leak tests were carried out in total.

4.2. Result and analysis

The apparent leak rates of the first level and second level batches, documented immediately after the dwell time, are shown in Fig. 6 in the semi-log scale. Batch 1 (and Batch 1-1) contained leaky packages while no leaky packages were included in Batch 1-2. Due to the same content of leak packages, Batch 1 and 1-1 produced practically the same leak rate curves, confirming the repeatability of the helium leak test conducted in this study. It also confirmed that the contribution of hermetic packages to leak rate signal is negligible and thus Eq. (4) is valid.

The initial apparent leak rate is a sum of the zero signal (pre-existing helium gas in the spectrometer chamber) and gas flowing out of leaky packages. The signal of Batch 1-2 was produced only by the zero signal and it was used to determine the zero signal time; it was determined to be 150 s after which the signal was stabilized to the ground level (10^{-10} atm-cc/s) [8].

The signal of Batch 1 of the measurement phase (after the zero signal time) is plotted in Fig. 7 in the linear scale. It was used for the subsequent regression analysis. The three criteria required for the forward-stepwise regression were first determined for the package and spectrometer used in the test. An optimal value of $e_i$ was obtained as $1 \times 10^{-6}$, which was comparable to the absolute regression error ($10^{-6.2}$ or 0.01%). The values of $e_{II}$ and $e_{III}$ were determined as $10^{-10}$ atm-cc/s and $5 \times 10^{-10}$ atm-cc/s, which correspond to the spectrometer sensitivity and the minimum measurable true leak rate, respectively.

The results of the regression analysis are presented in Table 1. Two leaky packages were identified in the batch. The true leak rates of the two packages were $5.90 \times 10^{-8}$ and $8.50 \times 10^{-7}$ atm-cc/s and the corresponding initial leak rates were $1.1 \times 10^{-6}$ and $3.92 \times 10^{-7}$ atm-cc/s. The regression results are compared with the experimental data in Fig. 6; they are virtually identical, indicating that the regression was effective. It is to be noted that discrete regression data points are shown in Fig. 7 in order to distinguish them from the experimental data.

4.3. Validation

As shown in Fig. 5, Batch 1-1 was broken into five sub-batches, among which Batch 1-1-2 and 1-1-5 were identified to have leaky packages. Individual tests were conducted for all packages within the two batches. Packages #19 and #41 were finally found to be
the above equation can be written as
allowable error norm in the regression analysis. The Taylor series of
the corresponding initial leak rates were 1.1 \times 10^{-2} and prediction. Considering that
deals with absolute (not logarithmic) difference between the data
of measurement phase are most important since the error norm
In the proposed regression analysis, data obtained at the early stage
1 \times 10^{-8} and 3.93 \times 10^{-3} atm-cc/s, for packages #19 and #41, respectively. The
The results of the regression are also plotted in Fig. 8.

ting section, which defines the applicable domain of the proposed
uncertainties. Although acceptable in practice, the cause of the dis-
crepancy is analyzed through an uncertainty analysis in the follow-
proved to define the extra constraints of the batch testing
for solution uniqueness.

5. Domain of application

As discussed earlier, the fine leak test has the limited applica-
tion domain due to the finite spectrometer sensitivity. The same
limitation should apply to the batch testing. In addition, the batch
testing data should be handled more carefully due to potential
multiple solutions (non-uniqueness) even when all leaky packages
in a batch lie in the detectable true leak range. The following anal-
ysis is provided to define the extra constraints of the batch testing
for solution uniqueness.

5.1. Error analysis

Let us consider a case where the combined signal of two leaky
packages (packages 1 and 2) is similar to a signal from a single lea-
ky package. It can be expressed as
\begin{equation}
\Omega \exp(-\lambda t) \approx \Omega_1 \exp(-\lambda_1 t) + \Omega_2 \exp(-\lambda_2 t) \quad \text{where} \quad \lambda = \frac{L}{VP_0}
\end{equation}

An “approximate-equal” sign is used in Eq. (8) to take account of the
allowable error norm in the regression analysis. The Taylor series of
the above equation can be written as
\begin{equation}
\Omega \left[1 - \frac{\lambda^2 t^2}{2} + \cdots \right] \approx \Omega_1 \left[1 - \lambda_1 t + \frac{\lambda_1^2 t^2}{2} + \cdots \right] + \Omega_2 \left[1 - \lambda_2 t + \frac{\lambda_2^2 t^2}{2} + \cdots \right]
\end{equation}

In the proposed regression analysis, data obtained at the early stage
of measurement phase are most important since the error norm
deals with absolute (not logarithmic) difference between the data
and prediction. Considering that \lambda t at the early stage is typically
on the order of 10^{-2}-10^{-3}, it is reasonable to assume that the con-
tribution of higher orders is negligible. Ignoring the higher order
terms, Eq. (9) yields
\begin{equation}
\Omega \approx \Omega_1 + \Omega_2, \quad \lambda \approx \lambda_1 + \lambda_2, \quad \lambda^2 \approx \lambda_1^2 + \lambda_2^2.
\end{equation}

The most ideal way to minimize the residual of regression is to
match the first two lowest order terms, which gives
\begin{equation}
\Omega = \Omega_1 + \Omega_2 \quad \text{and} \quad \lambda = \lambda_1 + \lambda_2
\end{equation}

Substituting Eq. (11) into the third equation of (10) yields
\begin{equation}
\Omega_1 \Omega_2 (\lambda_1 - \lambda_2)^2 \approx 0
\end{equation}

The left-hand side of the above equation can be a representation of the relative
residual of regression, a smaller value of which indicates higher possibility of false regression. Two possible cases to satisfy the equality are: (1) the true leak rates of two packages are identical, or (2) the product of two initial apparent leak rates is very small.

Fig. 9 illustrates the relative residual in Eq. (12) as a function of the true leak rate of the second package (L2) when L1 = 10^{-2} atm-
cc/s (the corresponding \lambda 1 is 4.64 \times 10^{-8}). The cavity volume and
test condition are identical to those used in the experiment. The
corresponding initial leak rates were calculated from Eq. (2). The

| Table 1 Regression results with batch and individual test data. |
|-----------------|------------------|
|                 | Package 1         | Package 2         |
| Batch L (L2) \times 10^{-7} | 59.0 (0.11)       | 8.50 (3.92)       |
| Package #19L (L2) \times 10^{-7} | 49.5 (0.11)       | 8.51 (3.93)       |
| Package #41L (L2) \times 10^{-7} |                   |                   |
| Difference (%) | 19.2 (0.0)        | 0.1 (0.3)         |

Fig. 8. Leak rate signals of a leaky package along with a batch that contains it: (a) Batch 1-1-2 and package #19 and (b) Batch 1-1-5 and package #41.
plot clearly shows two regimes where the regression is likely to fail; near $\lambda_1 \approx \lambda_2$ (or $\lambda_1 \approx L_1$) and too small $\Omega_1/\Omega_2$ (when $L_2$ is too small or too high).

### 5.2. Simulation

A series of batch tests was simulated to illustrate the above limitations. As before, the true leak rate of the first package ($L_1$) was fixed as $10^{-7}$ atm-cc/s while that of the second package ($L_2$) was changed from $L_1/100$ to $26L_1$. The corresponding initial leak rates were obtained from Eq. (2). The result of regression analyses is summarized in Table 2. The prediction failed at the two regions identified earlier; when two true leak rates were identical and when one of leaky packages had a very low initial leak rate close to the measurement sensitivity. In the other cases, the regression predicted leaky packages very accurately (better than the experimental case due to the absence of measurement uncertainties).

The regression analysis was extended to three leaky packages. Three representative cases are presented in Table 3. In the first case the three packages with true leak rates in the median range (Table 2) are considered; the result is accurate in not only the number of leaky packages but also the true leak rates. The second case, which contained two packages with low initial leak rates, produced the accurate number of leaky packages but incorrect true leak rates information for two packages with low initial leak rates. In the third case a package with the true leak rate of $2.6 \times 10^{-6}$ atm-cc/s, for which the regression failed to detect in the previous two-package simulation, replaced one of the packages. The regression failed again to capture this package.

This simulation study addresses clearly the application limit of the forward-stepwise regression analysis for the fine leak batch test; a feasible range of true leak rates is narrower than that of the individual test. In the simulations, the corresponding sensitivity limit of apparent leak rate was found to be approximately one order higher than the measurement sensitivity, i.e., $10^{-9}$ atm-cc/s. It is speculated that this value will be a little bit larger for real tests due to measurement uncertainties.

### 6. Discussion: advanced formulations

When the true leak rates of two leaky packages are nearly identical, the regression analysis may recognize them as a single package with a doubled initial leak rate. It can be avoided by implementing “caps” for the initial leak rate as a function of the true leak rate, i.e., $c(f(L_i) \leq \Omega_i < c(f(L_i))$. The caps should be set based on a proper correlation such as Eq. (2). For example, if the failed case in Table 2 due to the identical true leak rate is solved using a constraint based on Eq. (2), the new prediction yields two leaky packages with true leak rates of $1.069 \times 10^{-7}$ and $0.936 \times 10^{-7}$ atm-cc/s, which are within an error range of 7%.

The existing models are not able to provide accurate relationship between true and initial leak rates for the whole range of measurable true leak rates [9]. It can be easily confirmed by applying Eq. (2) for the two leaky test packages; it produces the initial leak rate as a function of the true leak rate and too small $\Omega_1/\Omega_2$ (when $L_2$ is too small or too high).

#### Table 3

Representative regression results for three leaky packages.

<table>
<thead>
<tr>
<th>Accurate prediction</th>
<th>1 (3.531)</th>
<th>0.2</th>
<th>5 (4.391)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_2 (\Omega_2) \times 10^{-7}$ (simulated)</td>
<td>1 (3.531)</td>
<td>0.2 (0.0)</td>
<td>25 (0.0209)</td>
</tr>
<tr>
<td>$L_1 (\Omega_1) \times 10^{-7}$ (predicted)</td>
<td>1.002 (3.531)</td>
<td>0.2004 (0.0207)</td>
<td>5.009 (4.391)</td>
</tr>
<tr>
<td>Difference (%)</td>
<td>0.02 (0.0)</td>
<td>0.02 (0.0)</td>
<td>0.20 (0.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accurate prediction</th>
<th>1 (3.531)</th>
<th>0.4</th>
<th>25 (0.0209)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_2 (\Omega_2) \times 10^{-7}$ (simulated)</td>
<td>1 (3.531)</td>
<td>0.04 (0.0651)</td>
<td>25 (0.0209)</td>
</tr>
<tr>
<td>$L_1 (\Omega_1) \times 10^{-7}$ (predicted)</td>
<td>1.109 (1.728)</td>
<td>0.308 (0.1531)</td>
<td>0.935 (1.722)</td>
</tr>
<tr>
<td>Difference (%)</td>
<td>0.10 (0.4)</td>
<td>0.963 (8139.2)</td>
<td>96.3 (8139.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Failed prediction</th>
<th>1 (3.531)</th>
<th>0.04</th>
<th>26 (0.0153)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_2 (\Omega_2) \times 10^{-7}$ (simulated)</td>
<td>1 (3.531)</td>
<td>0.04 (0.0651)</td>
<td>26 (0.0153)</td>
</tr>
<tr>
<td>$L_1 (\Omega_1) \times 10^{-7}$ (predicted)</td>
<td>1.022 (3.438)</td>
<td>0.290 (0.1634)</td>
<td>1.002 (7.062)</td>
</tr>
<tr>
<td>Difference (%)</td>
<td>0.02 (2.6)</td>
<td>151 (151)</td>
<td>151 (151)</td>
</tr>
</tbody>
</table>
where a batch with more than three leaky packages is split into as many sub-batches as needed until the number of leaky packages in all sub-batches becomes less than 3.

7. Conclusions

An effective way of analyzing the fine leak rate batch test was proposed to infer the information of leaky metal-sealed MEMS packages within a batch. The forward-stepwise regression analysis scheme was employed to make the quantitative inference possible. The scheme was implemented for wafer-level hermetic MEMS packages. The true leak rates of the two packages were obtained from a single batch test data. The results were validated subsequently by extra batch test data as well as individual tests. Errors associated with the regression analysis were discussed to define the applicable domain of the proposed scheme. Advanced formulations were also suggested to extend the domain for a wider range of applications.

References