Tolerance Stack Analysis Methods

Fritz Scholz*
Research and Technology
Boeing Information & Support Services

December 1995

Abstract

The purpose of this report is to describe various tolerance stacking methods without going into the theoretical details and derivations behind them. For those the reader is referred to Scholz (1995). For each method we present the assumptions and then give the tolerance stacking formulas. This will allow the user to make an informed choice among the many available methods.

The methods covered are: worst case or arithmetic tolerancing, simple statistical tolerancing or the RSS method, RSS methods with inflation factors which account for nonnormal distributions, tolerancing with mean shifts, where the latter are stacked arithmetically or statistically in different ways, depending on how one views the trade-off between part to part variation and mean shifts.

*Boeing Information & Support Services,
P.O. Box 3707, MS 7L-22, Seattle WA 98124-2207,
e-mail: fritz.scholz@grace.rt.cs.boeing.com

Glossary of Notation by Page of First Occurrence

term	**meaning**	**page**
σ, σ_i	standard deviation, describes spread of a statistical distribution for part to part variation	2, 15
L_i	actual value of i$^{\text{th}}$ detail part length dimension	4
G	gap, assembly criterion of interest, usually a function (sum) of detail dimensions	4
λ_i	nominal value of i$^{\text{th}}$ detail part dimension	4
T_i	tolerance value for i$^{\text{th}}$ detail part dimension	6
γ	nominal gap value, assembly criterion of interest	6
ϵ_i	difference between actual and mean (nominal) value of i$^{\text{th}}$ detail part dimension: $\epsilon_i = L_i - \lambda_i$ if mean $\mu_i =$ nominal λ_i, and $\epsilon_i = L_i - \mu_i$ if $\mu_i \neq \lambda_i$	6
a_i	coefficient for the i$^{\text{th}}$ term in the linear tolerance stack: $G = a_1 L_1 + \ldots + a_n L_n$, often we have $a_i = \pm 1$	7
X_i	actual value of i$^{\text{th}}$ input to sensitivity analysis; in length stacking X_i and L_i are equivalent	7
Y	output from sensitivity analysis; in length stacking Y and G are equivalent	7

Glossary of Notation by Page of First Occurrence

term	**meaning**	**page**
f	smooth function relating output to inputs in sensitivity analysis: $Y = f(X_1, \ldots, X_n)$ $Y = f(X_1, \ldots, X_n) \approx a_0 + a_1 X_1 + \ldots + a_n X_n$ $a_i = \partial f(\nu_1, \ldots, \nu_n)/\partial \nu_i, \qquad i = 1, \ldots, n$ $a_0 = f(\nu_1, \ldots, \nu_n) - a_1\nu_1 - \ldots - a_n\nu_n$	7
ν_i	nominal value of i^{th} input to sensitivity analysis in length stacking ν_i and λ_i are equivalent	8
ν	nominal output value from a sensitivity analysis in length stacking ν and γ are equivalent	8
T_{assy}	generic assembly tolerance derived by any method	9
$T_{\text{assy}}^{\text{arith}}$	assembly tolerance derived by arithmetic tolerance stacking (worst case method) $T_{\text{assy}}^{\text{arith}} = \lvert a_1 \rvert\, T_1 + \ldots + \lvert a_n \rvert\, T_n$	11
T_{detail}	tolerance common to all parts	11
ρ_i	tolerance ratio $\quad \rho_i = T_i/T_1$	11
$T_{\text{assy}}^{\text{stat}}$	assembly tolerance derived by statistical tolerance stacking (RSS method) $T_{\text{assy}}^{\text{stat}} = \sqrt{a_1^2 T_1^2 + \ldots + a_n^2 T_n^2}$	14

Glossary of Notation by Page of First Occurrence

term	meaning	page
$T_{\text{assy}}^{\text{stat}}(\text{Bender})$	assembly tolerance derived by statistical tolerance stacking (RSS method) using Bender's inflation factor of 1.5 $T_{\text{assy}}^{\text{stat}} = 1.5\sqrt{a_1^2T_1^2+\ldots+a_n^2T_n^2}$	16
c_i, $\boldsymbol{c}$, c	inflation factor for part variation distribution	17
$T_{\text{assy}}^{\text{stat}}(\boldsymbol{c})$	assembly tolerance derived by statistical tolerance stacking (RSS method) using distributional inflation factors $T_{\text{assy}}^{\text{stat}}(\boldsymbol{c}) = T_{\text{assy}}^{\text{stat}}(c_1,\ldots,c_n) = \sqrt{(c_1a_1T_1)^2+\ldots+(c_na_nT_n)^2}$	19
k	delimiter for the rectangular portion of the trapezoidal density	21
p	area of middle box of DIN-histogram density	23
g	half width of middle box of DIN-histogram density	23
μ_i	actual process mean for i^{th} detail part dimension	25
Δ_i	shift of process mean from nominal: $\Delta_i = \mu_i - \lambda_i$	25
η_i, $\boldsymbol{\eta}$	fraction of absolute mean shift in relation to T_i $\eta_i = \lvert\Delta_i\rvert/T_i$, $\boldsymbol{\eta} = (\eta_1,\ldots,\eta_n)$	25, 26

Glossary of Notation by Page of First Occurrence

term	meaning	page
L_i, U_i	lower and upper tolerance/specification limits: $L_i = \lambda_i - T_i$, $U_i = \lambda_i + T_i$	25
C_{pk}	a process capability index which accounts for mean shifts	25
$T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta})$	assembly tolerance derived by arithmetic stacking of mean shifts and RSS stacking of remaining normal variation; fixed T_i with tradeoff between mean shift and part variation $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}) = T_{\text{assy}}^{\Delta,\text{arith},1}(\eta_1, \ldots, \eta_n)$ $= \eta_1 \lvert a_1 \rvert T_1 + \ldots + \eta_n \lvert a_n \rvert T_n + \sqrt{[(1-\eta_1)a_1T_1]^2 + \ldots + [(1-\eta_n)a_nT_n]^2}$	26
T_i'	part tolerance based on part to part variation, either $T_i' = 3\sigma_i$ or $T_i' =$ half width of distribution interval	28, 31
$T_{\text{assy}}^{\Delta,\text{arith},2}(\boldsymbol{\eta})$	assembly tolerance derived by arithmetic stacking of mean shifts and RSS stacking of remaining normal variation; inflated T_i to accommodate mean shifts ($\leq \eta_i T_i$) under fixed $T_i' = 3\sigma_i = T_i/(1-\eta_i)$ part variation $T_{\text{assy}}^{\Delta,\text{arith},2}(\boldsymbol{\eta}) = T_{\text{assy}}^{\Delta,\text{arith},2}(\eta_1, \ldots, \eta_n)$ $= \eta_1 \lvert a_1 \rvert T_1'/(1-\eta_1) + \ldots + \eta_n \lvert a_n \rvert T_n'/(1-\eta_n) + \sqrt{(a_1T_1')^2 + \ldots + (a_nT_n')^2}$	28

Glossary of Notation by Page of First Occurrence

term	meaning	page
$T_{\text{assy}}^{\Delta,\text{arith}}(\boldsymbol{\eta}, \boldsymbol{c})$	assembly tolerance derived by statistical stacking (RSS method) using distributional inflation factors and arithmetic stacking of mean shifts $T_{\text{assy}}^{\Delta,\text{arith}}(\boldsymbol{\eta}, \boldsymbol{c}) = \eta_1 \vert a_1 \vert \, T_1 + \ldots + \eta_n \vert a_n \vert \, T_n$ $+\sqrt{[(1-\eta_1)c_1 a_1 T_1]^2 + \ldots + [(1-\eta_n)c_n a_n T_n]^2}$	29
σ_μ	standard deviation for mean shift distribution	33
$c_{\mu,i}$, $\boldsymbol{c}_\mu$, c_μ,	inflation factors for mean shift distributions	33, 33, 35
$T_{\text{assy}}^{\Delta,\text{stat},1}(\boldsymbol{\eta}, \boldsymbol{c}, \boldsymbol{c}_\mu)$	assembly tolerance derived by RSS stacking of mean shifts, RSS stacking of part variation and arithmetically stacking these two, assuming fixed part variation expressed through T_i' $T_{\text{assy}}^{\Delta,\text{stat},1}(\boldsymbol{\eta}, \boldsymbol{c}, \boldsymbol{c}_\mu) = \sqrt{c_1^2 a_1^2 T_1'^2 + \ldots + c_n^2 a_n^2 (1-\eta_n)^2 T_n'^2}$ $+\sqrt{c_{\mu,1}^2 a_1^2 \eta_1^2 T_1'^2/(1-\eta_1)^2 + \ldots + c_{\mu,n}^2 a_n^2 \eta_n^2 T_n'^2/(1-\eta_n)^2}$	34
R_i, $\boldsymbol{R}$	relative mean shift $\boldsymbol{R} = (R_1, \ldots, R_n)$ $R_i = \Delta_i/(\eta_i T_i), \qquad -1 \le R_i \le 1$	37
$\sigma(R_i)$	standard deviation of R_i	37

Glossary of Notation by Page of First Occurrence

term	**meaning**	**page**
w_i	a tolerance weight factor $w_i = a_i T_i / \sqrt{\sum_{j=1}^n a_j^2 T_j^2}\ , \qquad \sum_{i=1}^n w_i^2 = 1$	38
$F(\boldsymbol{R})$	inflation factor for given mean shift factor $\boldsymbol{R}$	38
$T_{\rm assy}^{\Delta,{\rm stat},2}(\eta)$	assembly tolerance derived by RSS stacking of mean shifts and RSS stacking of part variation which can increase with decrease in mean shifts; η is the common bound on all part mean shift fractions $T_{\rm assy}^{\Delta,{\rm stat},2}(\eta) = \left(\sqrt{1-\eta+\eta^2/2} + \eta\sqrt{3}\right) \times \sqrt{a_1^2 T_1^2 + \ldots + a_n^2 T_n^2}$	39
$T_{\rm assy}^{\Delta,{\rm arith,r}}(\boldsymbol{\eta}, \boldsymbol{c})$	reduced assembly tolerance using the factor .927 on the RSS part of $T_{\rm assy}^{\Delta,{\rm arith}}(\boldsymbol{\eta}, \boldsymbol{c})$ $T_{\rm assy}^{\Delta,{\rm arith,r}}(\boldsymbol{\eta}, \boldsymbol{c}) = \eta_1 \vert a_1 \vert\ T_1 + \ldots + \eta_n \vert a_n \vert\ T_n + .927\ \sqrt{[(1-\eta_1)c_1 a_1 T_1]^2 + \ldots + [(1-\eta_n)c_n a_n T_n]^2}$	41

1 Introduction and Overview

Tolerance stack analysis methods are described in various books and papers, see for example Gilson (1951), Mansoor (1963), Fortini (1967), Wade (1967), Evans (1975), Cox (1986), Greenwood and Chase (1987), Kirschling (1988), Bjørke (1989), Henzold (1995), and Nigam and Turner (1995). Unfortunately, the notation is often not standard and not uniform, making the understanding of the material at times difficult. For a critical review of these and some new methods and the mathematical derivation behind them see Scholz (1995).

Although the above cited references date back as far as Gilson's 1951 monograph, he provides several older references, namely Gramenz (1925), Ettinger and Bartky (1936), Rice (1944), Epstein (1946), Bates (1947, 1949), Nielson (1948), Gladman (1945), Loxham (1947) and some not associated with a person and thus omitted here. So far we have not been able to obtain any of these references, but it appears doubtful that anything beyond straight arithmetic or statistical tolerancing is contained in these. However, it would be of interest to find out who first proposed these two cornerstones of tolerancing and the various nuances that have followed.

There are no doubt many other sources which are internal to various companies and thus not very accessible to most people. For example, Wade (1967) mentions an article on statistical tolerancing by Backhaus and Fielden that appeared in an I.B.M. Corporation in-house publication. So far we have not been able to get a copy of this article. Other in-house writings on the subject are protected, such as Boeing's "proprietary" Tolerancing-Design Guide (1990) by Griess. Other companies have made their tolerancing guides widely available. As an example we cite the Motorola guide, authored by Harry and Stewart (1988). Unfortunately we have not been able to come up with sound, theoretical underpinnings for their proposed methods for dealing with mean shifts and thus we will omit them here. See Scholz (1995) for some discussion.

It is of interest to examine how the ASME Y14.5M-1994 standard and its companion ASME Y14.5.1M-1994 treat this subject. The former contains a very short Section 2.16, pp 38-39, which briefly mentions the basic forms of arithmetic and statistical tolerancing in connection with a new drawing symbol indicating a statistical tolerance, namely $\langle \mathrm{ST} \rangle$. This symbol is introduced there for the first time and it is to be expected that future editions of this standard will move toward taking advantage of statistical tolerance stacking. At this point the above symbol indicates that tolerances set with this symbol are to be monitored by statistical process control methods. How that is done is still left up to the user. Other symbols with similar intent are already in use in various companies.

Typically any exposition on tolerancing will include the two cornerstones, arithmetic and statistical tolerancing. We will make no exception, since these two methods provide conservative and optimistic benchmarks, respectively.

Under arithmetic tolerancing it is assumed that the detail part dimensions can have **any** value within the tolerance range and the arithmetically stacked tolerances describe the range of **all** possible variations for the assembly criterion of interest.

In the basic statistical tolerancing scheme it is assumed that detail part dimensions vary randomly according to a normal distribution, centered at the midpoint of the tolerance range and with its $\pm 3\sigma$ spread covering the tolerance interval. For given part dimension tolerances this kind of statistical analysis typically leads to much tighter assembly tolerances, or for given assembly tolerance it requires considerably less stringent tolerances for detail part dimensions, resulting in significant savings in cost or even making the difference between feasibility or infeasibility of a proposed design.

Practice has shown that the results are usually not quite as good as advertised. Assemblies often show more variation in the toleranced dimension than predicted by the statistical tolerancing method. The causes for this lie mainly in the violation of various distributional assumptions, but sometimes also in the misapplication of the method by not understanding the assumptions. Not wanting to give up on the intrinsic gains of the statistical tolerancing method one has tried to relax these distributional assumptions in a variety of ways. As a consequence such assumptions are more likely to be met in practice.

One such relaxation is to allow other than normal distributions. Such distributions essentially cover the tolerance interval with a wider spread, but are still centered on the tolerance interval midpoint. This results in somewhat less optimistic gains than those obtained under the normality assumption, but it usually still yields better results than those given by arithmetic tolerancing, especially for tolerance chains involving many detail parts.

Another relaxation of assumptions concerns the centering of the distribution on the tolerance interval midpoint. The realization that it is difficult to center any process exactly where one wants it to be has led to several mean shift models. In these the distribution may be centered anywhere within a certain small neighborhood around the nominal tolerance interval midpoint, but usually it is still assumed that the distribution is normal and its $\pm 3\sigma$ spread is within the tolerance limits. This means that while we allow some shift in the detail process mean we either require a simultaneous reduction in part variability or we have to widen the tolerance interval. The mean shifts are then stacked in worst case fashion. The variation of the shifted distributions is stacked statistically. The overall assembly tolerance then becomes (in worst case fashion) a sum of two parts, consisting of an arithmetically stacked mean shift contribution and a term reflecting the statistically stacked part variation distributions. It turns out that our cornerstones of arithmetic and statistical tolerancing are special cases of this more general model, which has been claimed (Greenwood and Chase, 1987) to unify matters.

However, there is another way of dealing with mean shifts which appears to be new, at least in the form presented here. It takes advantage of statistical stacking of mean shifts and stacking that aggregate in worst case fashion with the statistical stacking of the part variation distributions. A precursor to this can be found in Desmond's discussion of Mansoor's (1963) paper. However, there it was pointed out that it leads to optimistic results. We discuss the issues involved and present several variations on that theme.

Other fixes augment the statistical tolerancing method with blanket tolerance inflation factors with more or less compelling reasons. It turns out that one of the above mentioned mean shift treatments results in just such an inflation factor, with the size of the factor linked explicitly to the amount of tolerated mean shift.

When dealing with tolerance stacking under mean shifts one has to take special care in assessing the risk of nonassembly. Typically only one tail of the assembly stack distribution is significant when operating at one of the two worst possible assembly mean shifts. One can take advantage of this by reducing the assembly tolerance by some small amount. We indicate briefly how this is done but refer to Scholz (1995) for more details.

2 Notation and Problem Formulation

The tolerance stacking problem arises in the context of assemblies from interchangeable parts because of the inability to produce or join parts exactly according to nominal. Either the relevant part dimension varies around some nominal value from part to part or it is the act of assembly that leads to variation.

For example, as two parts are joined via matching hole pairs there is not only variation in the location of the holes relative to nominal centers on the parts but also the slippage variation of matching holes relative to each other when fastened.

Thus there is the possibility that the assembly of such interacting parts will not function or won't come together as planned. This can usually be judged by one or more assembly criteria, say $G_1, G_2, \ldots$.

Here we will be concerned with just one such assembly criterion, say G, which can be viewed as a function of the part dimensions $L_1, \ldots, L_n$. A simple example is illustrated in Figure 1, where $n = 6$ and

$$\begin{aligned} G &= L_1 - (L_2 + L_3 + L_4 + L_5 + L_6) \\ &= L_1 - L_2 - L_3 - L_4 - L_5 - L_6 \end{aligned} \tag{1}$$

is the clearance gap of interest. It determines whether the stack of cogwheels will fit within the case or not. Thus it is desired to have $G > 0$, but for functional performance reasons one may also want to limit G from above.

A graphical representation of equation (1) is given in Figure 2, where the various dimensions L_1, L_6, L_5, L_4, L_3, and L_2 are represented by vectors chained together, L_1 butting into L_6, L_6 butting into L_5 (after changing direction), L_5 butting into L_4, L_4 butting into L_3, and L_3 butting into L_2. The remaining gap to make L_2 butt up to L_1 is the assembly tolerance gap of interest, namely G. This type of linkage is called a tolerance path or tolerance chain. Note that the arrows point right for positive contributions and left for negative ones.

As was pointed out before, the actual lengths L_i may differ from the nominal lengths λ_i by some amount. If there is too much variation in the L_i there may well be significant problems in satisfying $G > 0$. Thus it is

Figure 1: Tolerance Stack Example

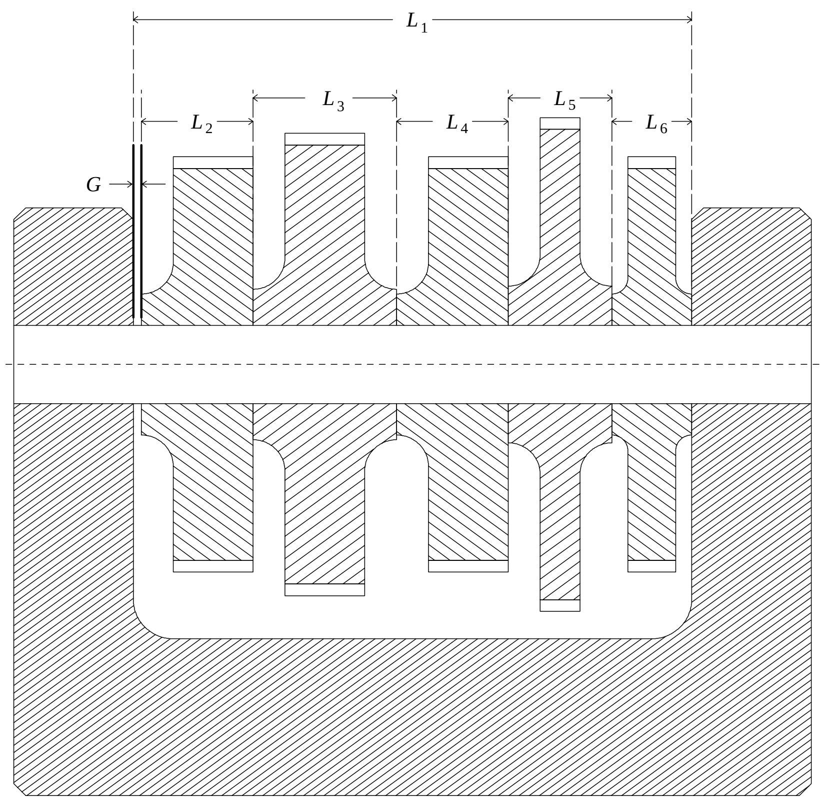
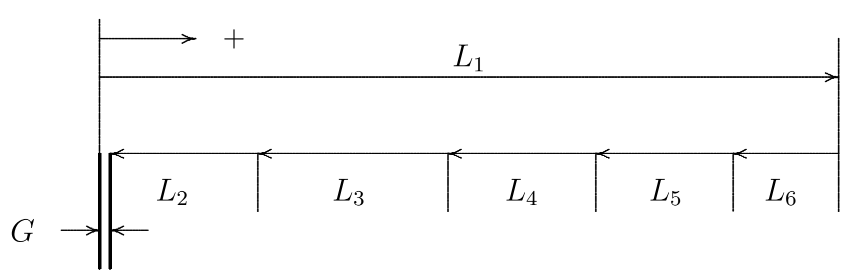


Figure 2: Tolerance Chain Graph



prudent to limit these variations through tolerances. Such tolerances, T_i, represent an "upper limit" on the absolute difference between actual and nominal values of the i^{th} detail part dimension, i.e., $|L_i - \lambda_i| \leq T_i$. It is mainly in the interpretation of this last inequality that the various methods of tolerance stacking differ.

The nominal value γ of G is usually found by replacing in equation (1) the actual L_i's by the corresponding nominal values λ_i, i.e.,

$$\gamma = \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 - \lambda_5 - \lambda_6 \ .$$

If the objective is to achieve a gap G that is positive and not too large (for other functional reasons) then one would presumably design the assembly in such a way that the nominal gap γ satisfies this goal, with the hope that the actual gap G be not too different from γ. Thus the quantity $G - \gamma$ is of considerable interest. It can be expressed as follows in terms of $\epsilon_i = L_i - \lambda_i$, the detail deviations from nominal,

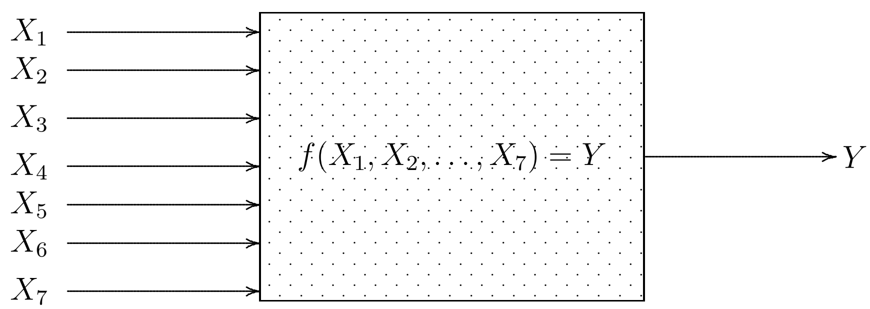
$$\begin{aligned} G - \gamma &= (L_1 - \lambda_1) - (L_2 - \lambda_2) - (L_3 - \lambda_3) \\ &\qquad\qquad - (L_4 - \lambda_4) - (L_5 - \lambda_5) - (L_6 - \lambda_6) \\ &= \epsilon_1 - \epsilon_2 - \epsilon_3 - \epsilon_4 - \epsilon_5 - \epsilon_6 \ . \end{aligned}$$

The main question of tolerance stacking is the bounding of the assembly error or assembly misfit $G - \gamma$ when given tolerance bounds T_i on the detail part errors, i.e. $|\epsilon_i| = |L_i - \lambda_i| \leq T_i$. In the following we will present several such bounds and state under what assumptions they are valid. Before doing so we generalize the above example to a generic tolerance chain and in the process widen the scope to smooth sensitivity analysis problems.

Above we had an assembly with a stack of six parts that involved one positive and five negative contributions. This can obviously be generalized to n detail parts with various configurations of positive and negative contributory directions in the tolerance chain. Hence in general we have:

$$G = a_1 L_1 + a_2 L_2 + a_3 L_3 + \ldots + a_n L_n \ ,$$

Figure 3: Function Box



where the coefficients $a_1, \ldots, a_n$ are either $+1$ or -1, independently of each other. Our introductory example had $n = 6$ and $a_1 = 1, a_2 = \ldots = a_6 = -1$. This then leads to

$$\begin{aligned} G - \gamma &= a_1(L_1 - \lambda_1) + a_2(L_2 - \lambda_2) + \ldots + a_n(L_n - \lambda_n) \\ &= a_1\epsilon_1 + a_2\epsilon_2 + \ldots + a_n\epsilon_n \end{aligned}$$

as the primary object of tolerance stack analysis.

From here it is only a small step to extending these methods to sensitivity analysis in general. Those not interested in this generalization can skip to the beginning of the next section.

Rather than butting parts end to end and forming an arithmetic sum of $\pm$ terms with some resultant output G, we can view this relation as a more general input/output relation. To get away from the restrictive notion of lengths we will use $X_1, \ldots, X_n$ as our inputs (in place of $L_1, \ldots, L_n$) and Y (in place of the gap G) as our output. However, here we allow more general rules of composition, namely

$$Y = f(X_1, \ldots, X_n) \ ,$$

where f is some known, smooth function which converts the inputs $X_1, \ldots, X_n$ into the output Y. This is graphically depicted in Figure 3. As an example of such a more general relationship consider some electronic device with components (capacitors, resistances, etc.) of various types. There may be several performance measures for such a device and Y may be any one of them. Given the performance ratings $X_1, \ldots, X_n$ of the various components, physical laws describe the output Y in some functional form, which typically is not linear. The design of such an electronic device is based on nominal values, $\nu_1, \ldots, \nu_n$, for the component ratings. However, the actual characteristics $X_1, \ldots, X_n$ will typically be slightly different from nominal, resulting in slight deviations for the actual $Y = f(X_1, \ldots, X_n)$ from the nominal $\nu = f(\nu_1, \ldots, \nu_n)$. Since these component deviations are usually small we can reduce this problem to the previous one of mechanically stacked parts by linearizing f, namely use

$$\begin{aligned} Y = f(X_1, \ldots, X_n) &\approx f(\nu_1, \ldots, \nu_n) + a_1(X_1 - \nu_1) + \ldots a_n(X_n - \nu_n) \\ &= a_0 + a_1 X_1 + \ldots + a_n X_n \end{aligned}$$

where $a_i = \partial f(\nu_1, \ldots, \nu_n)/\partial \nu_i$, $i = 1, \ldots, n$ and

$$a_0 = f(\nu_1, \ldots, \nu_n) - a_1\nu_1 - \ldots - a_n\nu_n \ .$$

Note: for this linearization to work we have to assume that f has continuous first partial derivatives at $(\nu_1, \ldots, \nu_n)$.

Aside from the term a_0 we have again the same type of arithmetic sum for our "assembly" criterion Y as we had in the mechanical tolerance stack. However, here the a_i are not restricted to the values ± 1. The additional term a_0[1] does not present a problem as far as variation analysis is concerned, since it is constant and known.

Again we like to understand how far Y may vary from the nominal $\nu = f(\nu_1, \ldots, \nu_n)$. From the above we have

$$\begin{aligned} Y - \nu &\approx a_1(X_1 - \nu_1) + \ldots a_n(X_n - \nu_n) \\ &= a_1\epsilon_1 + \ldots + a_n\epsilon_n \ , \end{aligned}$$

i.e., just as before, the only difference being that the a_i are not restricted to ± 1. Since all the tolerance stacking formulas to be presented below will be

[1] it is based on the nominal and known quantities $\nu_1, \ldots, \nu_n$

given in terms of these a_i and since nowhere use was made of $a_i = \pm 1$, it follows that they are valid for general a_i and thus for the sensitivity problem.

There are situations in which a functional relation $Y = f(X_1, \ldots, X_n)$, although smooth, is not very well approximated by a linear function, at least not over the range of variation envisioned for the X_i. In that case one could use a quadratic approximation to capture any relevant curvature in f. Tolerance stacking methods using this approach are covered in Cox (1986). These methods are fairly complex and still quite restrictive in the assumptions under which they are valid. Of course, it may be possible to extend these methods along the same lines as presented here for linear tolerance stacks.

As noted above, the linearization will work only for smooth functions f. To illustrate this with a counterexample, where linearization fails completely, consider the function

$$f(X_1, X_2) = \sqrt{X_1^2 + X_2^2}$$

which can be viewed as the distance of a hole center from the nominal origin $(0, 0)$. This function does not have derivatives at $(0, 0)$, its graph in 3-space looks like an upside cone with its tip at $(0, 0, 0)$. There can be no tangent plane at the tip of that cone and thus no linearization. Another example where such linearization fails is discussed in Altschul and Scholz (1994). It involves hinge mating and the problem arises due to simultaneous and thus minimum gap requirements.

In presenting the tolerance stacking formulas we will return to using L_i and λ_i for the part dimensions and nominals. Those that wish to apply these concepts to sensitivity analysis should have no problem replacing L_i, G, λ_i by X_i, Y, ν_i, respectively.

3 Tolerance Stacking Formulas

In this section we will present various formulas for tolerance stacking. By tolerance stacking we mean a rule that combines the detail tolerances T_i into an assembly tolerance T_{assy}. Typically T_{assy} is a monotone increasing function of the T_i. Thus, if the resulting T_{assy} is too large, one can counteract that by reducing all or some of the T_i, which usually makes for costlier part production. On the other hand, if T_{assy} is smaller than required for successful assembly fit, then one can loosen the detail tolerances T_i, with some possibility of cost reduction.

Why do we have more than one formula for tolerance stacking and why so many? One reason for this is that these methods have evolved and are still evolving, partly responding to economic pressures and partly because of the nature of the problem. Namely, it all depends on what assumptions one is willing to make.

Fewer assumptions entail broader applicability but one also will get less out of a tolerance stack analysis, i.e., one will wind up with fairly wide assembly tolerance limits or, when trying to counteract that through the T_i, with very tight and thus costly detail tolerance requirements.

With more knowledge about the manufacturing processes one may feel comfortable with more assumptions, resulting in tighter assembly tolerance limits or, if those can be relaxed, with looser detail tolerance requirements.

Thus it is very important to be aware of the assumptions behind the various methods. We will begin the presentation of stacking methods with the worst case or arithmetic method, which tends to be most conservative. This is followed by the conventional RSS or statistical tolerance stacking method, which tends to be on the optimistic side. This results from imposing some rather stringent assumptions. If the arithmetic stacking method gives satisfactory assembly tolerance results, then there is little motivation to try any of the other methods, except possibly to relax detail tolerances to achieve cost reduction. If the RSS method does not give satisfactory assembly tolerance results, then any of the other methods will not make matters any better. Then the only recourse is to tighten detail tolerances or, if that is not feasible, change the design.

After discussing these two basic and well known methods we will discuss several hybrid tolerance stacking methods which impose assumptions which are more likely to be met in practice. As a result the assembly tolerances lie somewhere between those corresponding to the two classical methods.

3.1 Arithmetic or Worst Case Tolerance Stacking

Assuming $|\epsilon_i| = |L_i - \lambda_i| \le T_i$ for all $i = 1, 2, \ldots, n$ we can bound $|G - \gamma|$ by

$$\boxed{T_{\rm assy}^{\rm arith} = |a_1|T_1 + |a_2|T_2 + \ldots + |a_n|T_n \; . \qquad (2)}$$

If $|a_i| = 1$ for all i, this simplifies to

$$T_{\rm assy}^{\rm arith} = T_1 + T_2 + \ldots + T_n \; .$$

The validity hinges solely on the above assumption. Thus, no matter how the detail dimensions L_i deviate from their nominal values λ_i within the constraint $|L_i - \lambda_i| \le T_i$, the difference $|G - \gamma|$ is guaranteed to be bounded by $T_{\rm assy}^{\rm arith}$. This guarantee is the main strength of this method. However, one should not neglect to make sure that the assumptions are met, i.e., detail parts need to be inspected to see whether $|L_i - \lambda_i| \le T_i$ or not.

The main weakness of the method is that $T_{\rm assy}^{\rm arith}$ grows more or less linearly with n. This is most easily seen when the detail part tolerance contributions $|a_i|T_i$ are all the same, i.e., $|a_i|T_i \equiv T_{\rm detail}$ in which case

$$T_{\rm assy}^{\rm arith} = n \cdot T_{\rm detail} \; .$$

By inverting this we get

$$T_{\rm detail} = \frac{T_{\rm assy}^{\rm arith}}{n} \; ,$$

which tells us how to specify detail tolerances from assembly tolerances. As assemblies grow, i.e., as n gets large, these requirements on the detail tolerances can get quite severe.

The linear growth of $T_{\rm assy}^{\rm arith}$ results from assuming a devil's advocate position in deriving the formula for $T_{\rm assy}^{\rm arith}$, namely by always taking the detail part variation in such a way as to make the assembly criterion G differ as much as possible from γ. This is the reason for the method's alternate name: worst case tolerancing.

If the detail tolerances are not all the same, it is more complicated to arrive at appropriate detail tolerances satisfying a given assembly tolerance requirement. For example, suppose $T_i = \rho_i T_1$ for $i = 2, \ldots, n$. Then

$$T_{\rm assy}^{\rm arith} = T_1 + \rho_2 T_1 + \ldots + \rho_n T_1 = T_1(1 + \rho_2 + \ldots + \rho_n)$$

so that

$$T_1 = \frac{T_{\text{assy}}^{\text{arith}}}{1 + \rho_2 + \ldots + \rho_n}, \qquad \text{and} \qquad T_i = \frac{\rho_i T_{\text{assy}}^{\text{arith}}}{1 + \rho_2 + \ldots + \rho_n}$$

for $i = 2, \ldots, n$. Thus relaxing or tightening $T_{\text{assy}}^{\text{arith}}$ by some factor affects all detail tolerances T_i by the same factor.

One may also want to treat the detail tolerances T_i in a more differentiated manner, i.e., leave some as they are and reduce other significantly in order to achieve the desired assembly tolerance. This easily done in iterative fashion using the forward formula (2).

The above considerations on how to set detail tolerances based on assembly tolerance requirements can be carried out for the other types of tolerance stacking as well and we leave it up to the reader to similarly use the various tolerance stacking formulas in reverse.

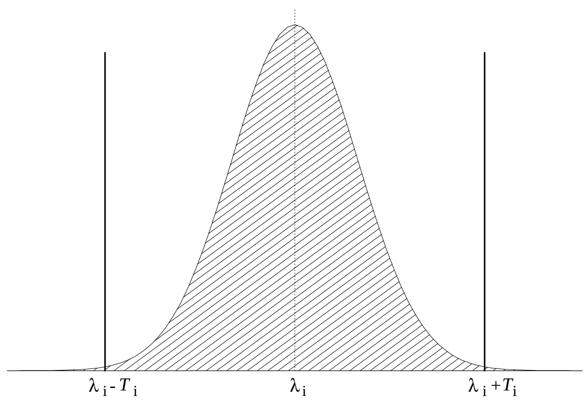
3.2 RSS Method or Statistical Tolerancing

Under this method of tolerance stacking a very important new element is added to the assumptions, namely that the detail variations from nominal are random and independent from part to part. In some sense this is a reaction to the worst case paradigm of the previous section which many feel is overly conservative. It is costly in the sense that it often mandates very tight detail tolerances.

That all deviations from nominal should arrange themselves in worst case fashion to yield the most extreme assembly tolerance is a rather **unlikely** proposition. However, it had the benefit of guaranteeing the resulting assembly tolerance. Statistical tolerancing in its classical form operates under two basic assumptions:

Centered Normal Distribution: Rather than assuming that the L_i can fall anywhere within the tolerance interval $[\lambda_i - T_i, \lambda_i + T_i]$, even to the point that someone maliciously and deliberately selects parts for worst case assemblies, we assume here that the L_i are normal random variables, i.e., vary randomly according to a **normal** distribution, **centered** on that same interval and with a $\pm 3\sigma$ spread equal to the span

Figure 4: Normal Distribution Over Tolerance Interval



of that interval, so that 99.73% of all L_i values fall within this interval, see Figure 4. The nature of the normal distribution is such that the L_i occur with higher frequency in the middle near λ_i and with less frequency near the interval endpoints. The match of the $\pm 3\sigma$ spread with the span of the detail tolerance span is supposed to express that almost all parts will satisfy the detail tolerance limits.

Deviations from nominal are not a deliberate act but inadvertant and due to forces not under our control. If these forces are several and influence the final deviation from the nominal value in independent fashion, then there are theoretical reasons (the central limit theorem of probability theory) supporting a normal distribution for L_i. However, it may not always be reasonable to assume that this normal distribution is exactly centered on the nominal value. This objection is the starting point for some of the hybrids to be discussed later.

Independent Detail Variation: The independence assumption is probably the most essential cornerstone of statistical tolerancing. It allows for some cancellation of variation from nominal.

Treating the L_i as random variables, we also demand that these random variables are (statistically) independent. This roughly means that the deviation $L_i - \lambda_i$ has nothing to do with the deviation $L_j - \lambda_j$ for $i \neq j$. In particular, the deviations will not be mostly positive or mostly negative. Under independence we expect to get a mixed bag of negative and positive deviations of various sizes which essentially leads to some variation cancellation in the adding process. Randomness alone does not guarantee such cancellation, especially not when all part dimension show random variation in the same direction. This latter phenomenon is exactly what the independence assumption intends to exclude.

Typically the independence assumption is reasonable when part dimensions pertain to different manufacturing/machining processes. However, situations can arise where this assumption is questionable. For example, several similar/same parts (coming from the same process) could be used in the same assembly. If this process is affected by a mean shift, then this mean shift will accumulate in worst case fashion for all parts coming from that process. Thermal expansion also tends to affect different parts similarly.

Under the above assumptions of centered normality and independence we can give the following statistical tolerance stacking formula

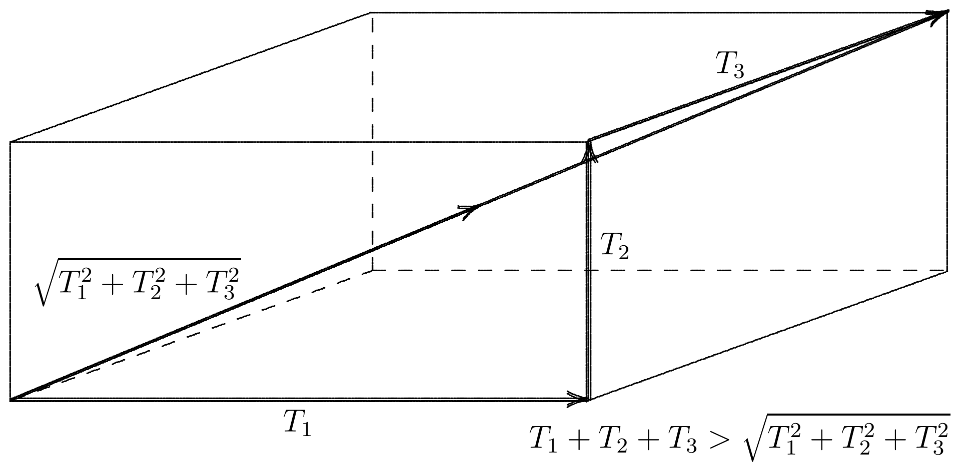
$$\boxed{\begin{aligned} T_{\text{assy}}^{\text{stat}} &= \sqrt{a_1^2 T_1^2 + a_2^2 T_2^2 + \ldots + a_n^2 T_n^2} \qquad (3) \\ &= \sqrt{T_1^2 + T_2^2 + \ldots + T_n^2} \end{aligned}}$$

where the latter formulation holds when $a_i = \pm 1$ for all $i = 1, \ldots, n$. The term RSS for this type of stacking stems from its abbreviation for *Root Sum of Squares*.

Typically $T_{\text{assy}}^{\text{stat}}$ is significantly smaller than $T_{\text{assy}}^{\text{arith}}$. For $n = 3$ the magnitude of this difference is easily visualized and appreciated by a rectangular box with side lengths T_1, T_2 and T_3. To get from one corner of the box to the diagonally opposite corner, one can traverse the distance $\sqrt{T_1^2 + T_2^2 + T_3^2}$ along that diagonal or one can go the long way and follow the three edges with lengths T_1, T_2, and T_3 for a total length $T_{\text{assy}}^{\text{arith}} = T_1 + T_2 + T_3$ as in Figure 5.

This reduction in assembly tolerance comes at a small price. Whereas $T_{\text{assy}}^{\text{arith}}$ bounds the assembly deviation $|G - \gamma|$ with absolute certainty, the

Figure 5: Pythagorean Shortcut



statistical tolerance stack $T_{\text{assy}}^{\text{stat}}$ bounds $|G-\gamma|$ only with some high assurance, namely with .9973 probability. The crookedness of .9973 results from the fact that the variation of G around γ is again normal[2] and that $\pm T_{\text{assy}}^{\text{stat}}$ represents a $\pm 3\sigma$ range for that variation. The 3 in 3σ is a nice round number, but the probability content (.9973) associated with it is not. One cannot have it both ways.

The small price, going from absolute certainty down to 99.73%, is not all. Recall that normal part variation, centered on the tolerance interval with $T_i \equiv 3\sigma_i$, and independence of variation from part to part are assumed as well.

3.3 RSS Method With Inflation Factors

Practice has shown that arithmetic tolerancing tends to give overly conservative results and that the RSS method is too optimistic, i.e., is not living up to the proclaimed 99.73% assembly fit rate. This means that actual assembly stack variations are wider than indicated by the $\gamma \pm T_{\text{assy}}^{\text{stat}}$ range. The reasons

[2]being a sum of independent, normally distributed length dimensions, without appeal to the central limit theorem

for this have been examined from various angles. We list here

independence: An important aspect of statistical tolerance stacking is the independence of variations from nominal for the detail parts participating in an assembly.

$3\boldsymbol{\sigma}_i = \boldsymbol{T}_i$: Does the $\pm T_i$ range really represent most or all of the detail part variation?

normality: Is the detail part variation reasonably represented by the normal distribution?

centered process: Is the process of part variation centered on the nominal, the midpoint of the tolerance interval?

One reason for a reduction in the efficacy of statistical tolerance stacking could be that the independence assumption is violated. We will not dwell on that issue too much except for some very specific modes of dependence such as random mean shifts or tooling errors. Dependence can take so many forms that it is difficult to cope with it in any systematic way. However, we will return to this later when we discuss mean shifts that are random.

One other possible reason for the optimism of the RSS method is that one basic premise, namely $T_i \equiv 3\sigma_i$, is not fulfilled. This can come about when manufacturing process owners, asked for the kind of tolerances they can hold, sometimes will respond with a $\pm T_i$ value which corresponds to a $\pm 2\sigma_i$ range. Reasons for this could be limited exposure to actual data. Values outside the $\pm 2\sigma_i$ range are hardly ever experienced[3] and if they do occur they may be rationalized away as an abnormality and then disappear from the conscious record. Thus, if T_i is specified with the misconception $T_i \equiv 2\sigma_i$, then T_i is too small by a factor 1.5. To correct for this, Bender (1962) suggests to multiply the $\pm T_{\text{assy}}^{\text{stat}}$ value by 1.5 and calls this process "benderizing," i.e.,

$$\boxed{T_{\text{assy}}^{\text{stat}}(\text{Bender}) = 1.5\,\sqrt{a_1^2 T_1^2 + \ldots + a_n^2 T_n^2} = 1.5\;\; T_{\text{assy}}^{\text{stat}}\,. \qquad (4)}$$

[3] The $\pm 2\sigma_i$ range contains about 95% of all variation under a normal curve.

The assumptions behind this formula are the same as those for (3) except that detail part tolerances correspond to $\pm 2\sigma_i$ rather than $\pm 3\sigma_i$ normal variation ranges.

This inflation factor 1.5 gives up a fair amount of the gain in $T_{\text{assy}}^{\text{stat}}$. In fact, for $n = 2$ it is more conservative than arithmetic tolerance stacking, since

$$1.5 \sqrt{a_1^2 T_1^2 + a_2^2 T_2^2} \geq |a_1| T_1 + |a_2| T_2 \ .$$

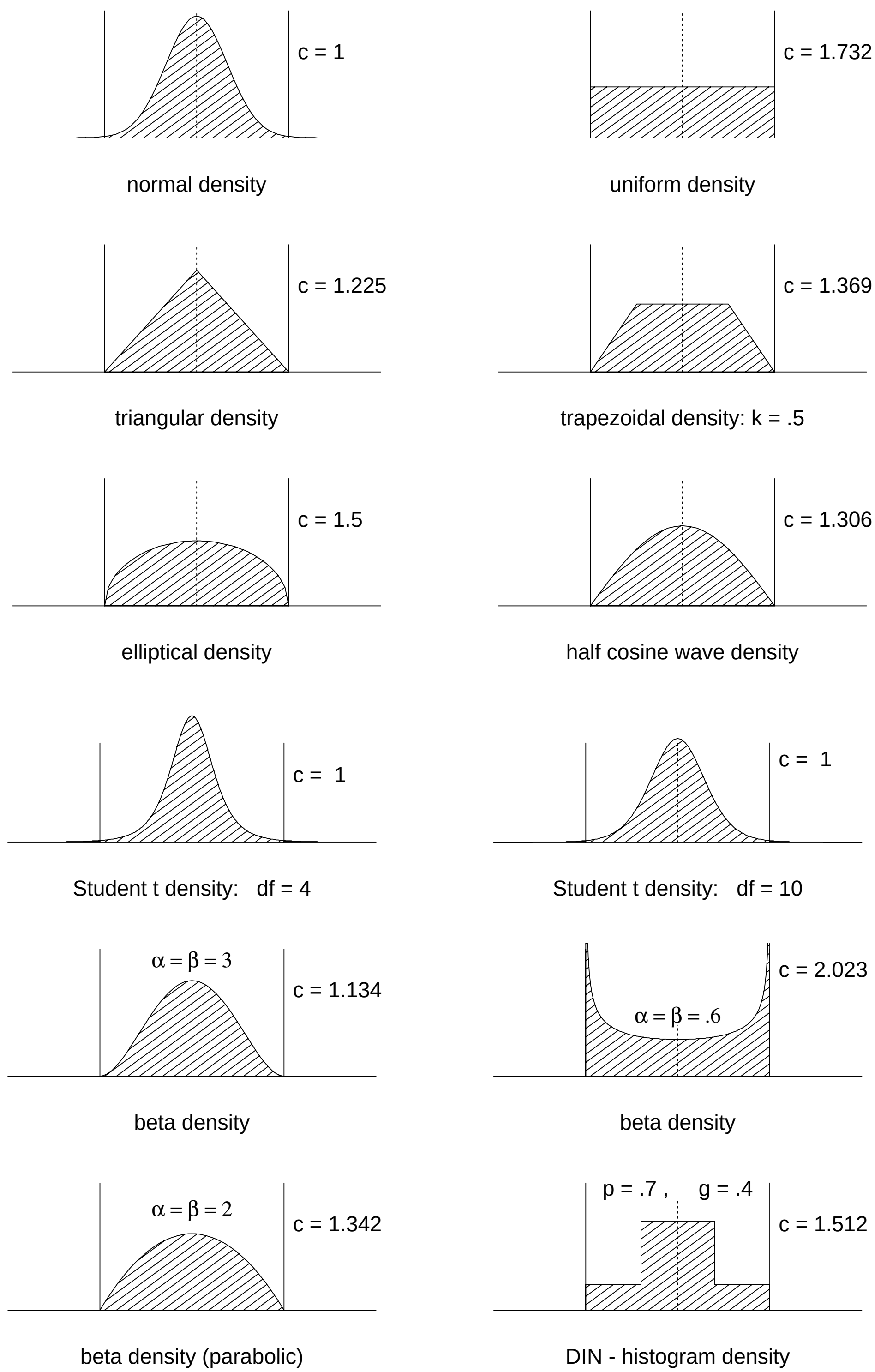
Of course, some may say that we should have used a 1.5 factor on the right side as well, because those tolerances are also misinterpreted. The rationale for the inflation factor is not altogether satisfactory, since it is based on ignorance and suppositions about meanings of T_i. What we have here is mainly a communications breakdown. If we do not have data about the part process capabilities, any tolerance analysis will stand on weak legs. If we have only limited data, then it should still be possible to avoid the mixup of $2\sigma_i$ with $3\sigma_i$ variation ranges. In fact, upper confidence bounds on $3\sigma_i$, based on limited data, will be quite conservative and thus should lead to conservative values $T_{\text{assy}}^{\text{stat}}$ when using such confidence bounds for T_i.

Although the normality assumption is well supported by the central limit theorem[4], there are processes producing detail part dimensions which are not normally distributed. Some such processes come about through tool wear, where part dimensions may start out at one end of the tolerance range and, as the tool wears, eventually wind up at the other end. The collection of such parts would then exhibit a more uniform distribution over the tolerance range.

Some people have simply postulated a somewhat wider distribution over the $\pm T_i$ tolerance range mainly for the purpose of obtaining an inflation factor to the RSS formula, see Gilson (1951), Mansoor (1963), Fortini (1967), Kirschling (1988), Bjørke (1989), and Henzold (1995). Several such distributions are illustrated in Figure 6 with the corresponding inflation factors c. Of course, one may find that different detail part variations warrant different inflation factors. Using such inflation factors $\boldsymbol{c} = (c_1, \ldots, c_n)$ for the n detail parts leads to the following modified statistical tolerance stacking formula:

[4]in the sense that a total detail part variation, made up more or less additively of many small random contributions, is approximately normal

Figure 6: Distribution Inflation Factors



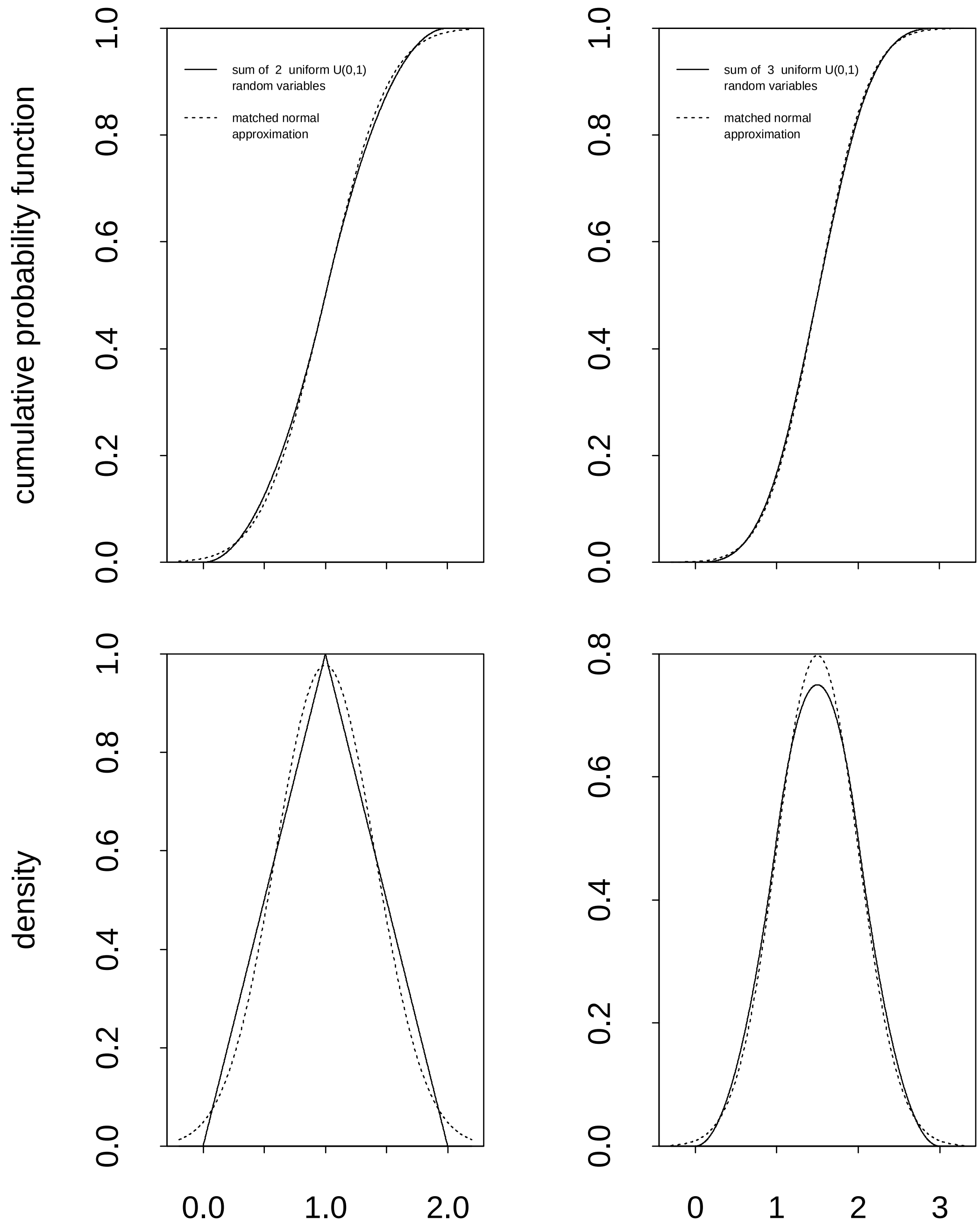
$$
\begin{aligned}
T_{\text{assy}}^{\text{stat}}(\boldsymbol{c}) &= T_{\text{assy}}^{\text{stat}}(c_1, \ldots, c_n) \\
&= \sqrt{(c_1 a_1 T_1)^2 + (c_2 a_2 T_2)^2 + \ldots + (c_n a_n T_n)^2}
\end{aligned} \tag{5}
$$

The underlying assumptions are that the part variations are independent and are characterized by possibly diverse distributions centered on the part tolerance intervals. These distributions, not necessarily normal, mostly cover the respective part tolerance intervals, either completely or by their $\pm 3\sigma_i$ ranges, see Figure 6.

The interpretation of this assembly tolerance stack is as before. Namely, one can expect that 99.73% of all assembly G gap values fall within $\gamma \pm T_{\text{assy}}^{\text{stat}}(\boldsymbol{c})$. Although the individual contributors to the stack may no longer be normally distributed we can still appeal to the central limit theorem to conclude that G is approximately normally distributed. Since the word "limit" in central limit theorem implies that the number of terms being added should be at least moderately large, it is worth noting that in many situations one can get fairly reasonable normal approximations already for $n = 2$ or $n = 3$ stacking terms.

One notable problem case among the distributions featured in Figure 6 is the uniform distribution. In that case the sum of two uniformly distributed random variables will in general have a trapezoidal density, which on the face of it cannot qualify as being approximately normal. If the two uniform distributions have the same width then this trapezoidal density becomes triangular. See the left side of Figure 7 where the top panel gives the cumulative distribution and its normal approximation and the bottom panel shows the corresponding densities for the sum of two random variables, uniformly distributed over the interval $(0, 1)$. The right side of Figure 7 shows the analogous comparisons for the sum of three such uniform random variables. Although the density comparison shows strong discrepancies for the sum of two uniform random terms, there appears to be much less difference for the cumulative distribution, since the undulating errors, visible for the densities, cancel out in the probability accumulation process. Thus the central limit theorem could be appealed to even in that case, if one is content

Figure 7: Central Limit Theorem Effect: small n



with somewhat rougher probability approximations

Note also that the normal approximation spreads out further than the approximated distribution. This would result in conservative assembly risk assessment. Rather than .27% of assemblies falling out of tolerance (under the normal approximation) it would be actually less under uniform detail part variation.

Before using inflation factors based on specific distributions one should make sure that such distributions are really more appropriate than the customary normal distribution. Such judgments should be based on data. If one has such validated concerns they may affect just one or two such contributors in (5) and leaving most other c factors equal to one. Note that c factors larger than one increase the assembly tolerance stack.

We view formula (5) mainly as a useful extension to formula (3) for just such situations where normality does not hold for all detail part dimensions. This way the behavior of one part process alone will not preclude us from performing a valid statistical tolerance analysis.

If one uses such distributions solely for getting some sort of inflation or protection factor without having any other justification, one should drop that distribution pretense and just admit to using an inflation factor for just such protection purpose.

Some of the distributions portrayed in Figure 6 require some comments or explanation. The uniform distribution can in some sense be viewed as a most conservative description of variation over a fixed interval. Among all symmetric, unimodal[5] distributions over such an interval it has the most spread or the largest standard deviation σ_i.

The trapezoidal density is uniform on the interval $[\lambda_i - kT_i, \lambda_i + kT_i]$, where k is some number in $[0, 1]$, and the density falls off linearly to zero over $[\lambda_i + kT_i, \lambda_i + T_i]$ and $[\lambda_i - T_i, \lambda_i - kT_i]$. The uniform and triangular density are special cases of the trapezoidal one.

[5]A density $f(x)$ is symmetric and unimodal about λ_i if $f(x)$ has same values for $x = \lambda_i \pm \Delta$ and if these values are nonincreasing as Δ moves away from zero. The beta density with $\alpha = \beta = .6$ is the only one in Figure 6 which is not unimodal. It is bimodal, since it has two separate peaks.

The elliptical density[6] consists of half an ellipse and is characterized by the requirement that one axis of the ellipse straddles the interval $\lambda_i \pm T_i$ and its other half axis has length $2/(\pi T_i)$.

Aside from the normal distribution the Student t-density is the only one among the illustrated distributions which has an unbounded range. This raises the issue of how to match up the range of such distributions with the finite range $[\lambda_i - T_i, \lambda_i + T_i]$. In the normal case it has been traditional to take $T_i \equiv 3\sigma_i$ with the normal distribution centered on λ_i. With that identification 99.73% of all detail parts of type i will vary within $[\lambda_i - T_i, \lambda_i + T_i]$. In the case of the Student t-distribution we have two options. We can either scale the t-distribution to match the probability content of .9973 over $[\lambda_i - T_i, \lambda_i + T_i]$ or we can again let $T_i \equiv 3\sigma_i$. In the former approach we will wind up with c-factors that are less than one, because each σ_i would typically be much smaller than $T_i/3$. The trouble with this approach is that with limited data it is very difficult to establish that $[\lambda_i - T_i, \lambda_i + T_i]$ captures 99.73% of all detail part dimensions.

The other approch, namely $T_i \equiv 3\sigma_i$, is much easier to implement with limited data and it leads to a c-factor which is one. The ease derives from the fact that standard deviations can be estimated with fairly limited data. However, the smaller the data set, the less certain we can be about the standard deviation estimate.

One detraction with using $T_i \equiv 3\sigma_i$ is that we will tend to see more detail parts out of tolerance. In using statistical tolerancing ideas there is no need to guarantee that all detail parts are within tolerance as is required under arithmetic tolerancing. In statistical tolerancing we only need to control the amount of part variation. Occasional detail parts which fall out of tolerance do not need to be sorted out. They actually may average out just fine in the assembly. Note that the two t-distributions illustrated in Figure 6 have different degrees of freedom and thus different detail fall-out rate.

The beta density comprises a rich family of shapes and for its mathematical form we refer to Scholz (1995). Here we only considered symmetric beta densities with paramters $\alpha = \beta$ and standard deviation $\sigma_i = T_i/\sqrt{2\alpha + 1}$.

[6]by some also called semicircular density, since the elliptical shape is the result of normalizing the total area under the density to one

Table 1: Distributional Inflation Factors

normal	1	uniform	1.732
triangular	1.225	trapezoidal	$\sqrt{3(1+k^2)/2}$
cosine half wave	1.306	elliptical	1.5
Student-t	1		
beta (symmetric)	$3/\sqrt{2\alpha+1}$		
histogram density (DIN)	$\sqrt{3}\,\sqrt{(1-p)(1+g)+g^2}$		

Finally, the box shaped or DIN-histogram density given in the lower right panel of Figure 6 is characterized by the two parameters (p, g), where p stands for the probability content or area of the middle box sitting on the interval $[\lambda_i - gT_i, \lambda_i + gT_i]$. The rest of the probability $(1-p)/2$ is distributed uniformly and in equal parts over the two other boxes filling up the rest of the interval $[\lambda_i - T_i, \lambda_i + T_i]$. This density figures prominently in the German attempts to standardize tolerance calculations, see Henzold (1995). However, these proposals still appear to be preliminary. The appealing aspect of this formulation of nonnormal variation is that one specifies two tolerance zones, the outside one, given by $\lambda_i \pm T_i$ and containing all variation, and the inside one, given by $\lambda_i \pm gT_i$ and containing usually most or $100p\%$ of all variation. The assumption of uniformity within each range is in some sense a most conservative stance, as pointed out above in connection with the uniform distribution.

Table 1 gives the formulas or numerical expressions for the c-factors of the distributions illustrated in Figure 6.

3.4 Normal RSS With Arithmetically Stacked Mean Shifts

One crucial assumption in the RSS method is that the normal distribution characterizing the variation of the detail parts over tolerance ranges be centered on the midpoint of such tolerance ranges. This may seem reasonable in that such midpoints represent detail part nominal values and it would seem plausible that most manufacturing process would aim for such nominals. Any deviations from such aim should balance equally so that the processes would then be centered on their respective nominals.

Unfortunately this may be wishful thinking. The manufacturing process owner, when presented with the tolerance range $\lambda_i \pm T_i$ may not necessarily set up the process with aim at λ_i. This could be for a variety of reasons. For example, the variability of the process may be considerably smaller than indicated by the given $\pm T_i$ tolerance range. In that case the process owner may not have the motivation in spending much effort on centering the process as long as the part variation stays within the tolerance interval. Furthermore, the owner could try to take advantage of this wider tolerance interval by moving the process off-center and thereby decreasing some other cost aspect, e.g., cost of labor, material, etc. Of course such strategy does not work when the process variability barely fits within the allotted $\lambda_i \pm T_i$.

Another reason for expecting some amount of mean shift, i.e., the process mean being shifted from the nominal center λ_i, is that even if one tries hard to set up a process with mean centered on some particular value, one is never fully successful. One will be off by some amount. A correction will suffer from the same weakness and if one corrects too often, one only adds variability to the process and thus making it only harder to discern mean shifts. Some may view this as a solution, but it seems to be a questionable one.

The main aspect of mean shifts is that they are a systematic component of detail part deviation from nominal. Their effect on assembly variation is the same for assembly after assembly. Thus it is especially important to control the possible negative impacts of such effects.

This is a good place to consider another aspect of assembly, namely that of tooling. Tools are used to aid in putting an assembly together. These are fixtures, jigs, or other devices that hold parts in place or are used in positioning parts to be fastened to each other. Part of the assembly process variation stems from the tool. Some of this variation happens anew each time an assembly is made. For example, the error of positioning a part relative to an index point is of this type. Such variation contributors should be handled as though they constitute a "part" with part to part variation. Often however, there are features of the tool that are more or less permanent, i.e., do not change from assembly to assembly. Examples are stops or index points. Although one wishes such tool features to coincide exactly with nominal positions, there are deviations from nominal which could be considered as mean shifts in the positioning process.

In considering mean shifts as a possibility one has to settle on how much mean shift one is willing to accept in a tolerance analysis. Of course, one also has to make sure that such mean shift assumptions are reasonable in practice, i.e., one will have to resort to statistical quality control methods to monitor the relevant processes for these assumptions. To quantify the amount of mean shift we introduce some notation. Let μ_i denote the process mean for the i$^{\text{th}}$ detail part dimension and let $\Delta_i = \mu_i - \lambda_i$ be the corresponding mean shift, i.e., the difference between process mean and nominal. It is useful to characterize the absolute difference in relation to the tolerance T_i, namely, $|\Delta_i| = \eta_i T_i$. Here the fraction $\eta_i = |\Delta_i|/T_i$ will typically be a number between 0 and 1. Usually, values $\eta_i = .10, .20$, or $.30$ will be most common. By focussing on the ratios η_i we want to control the mean shifts in relation to the tolerances T_i, i.e., large tolerances usually permit also larger mean shifts and it will be more reasonable to assume a common value for all the η_i.

While allowing some amount of mean shift we will however insist that the total process variation will still be contained in the tolerance interval $[\lambda_i - T_i, \lambda_i + T_i] = [L_i, U_i]$, i.e., be within the upper and lower part specification/tolerance limits, and that the process variation be normal. In terms of the process capability index C_{pk} this means

$$C_{pk} = \min\left(\frac{U_i - \mu_i}{3\sigma_i}, \frac{\mu_i - L_i}{3\sigma_i}\right) = \frac{T_i - |\Delta_i|}{3\sigma_i} = \frac{T_i - \eta_i T_i}{3\sigma_i} = \frac{(1-\eta_i)T_i}{3\sigma_i} \geq 1 \ .$$

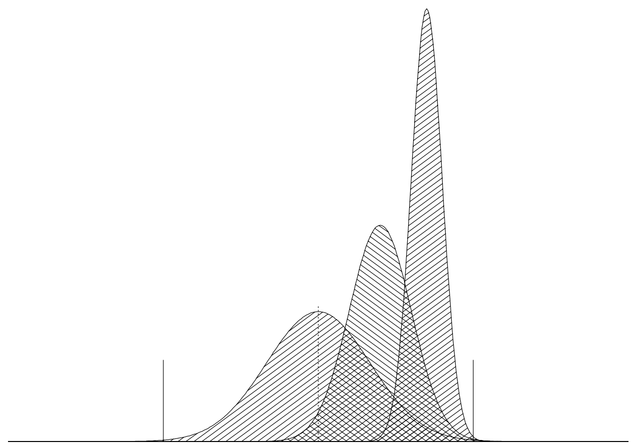
Assuming the highest amount of variability within these constraints, i.e., $C_{pk} = 1$, we have $3\sigma_i = (1-\eta_i)T_i$ and we see that increasing the mean shift ratio η_i while holding T_i fixed entails that the process variability σ_i be decreased, see Figure 8. This is also seen clearly from

$$T_i = (1-\eta_i)T_i + \eta_i T_i = 3\sigma_i + |\Delta_i|$$

where an increase in $|\Delta_i|$ needs to be traded off against a decrease in σ_i in order to maintain a fixed T_i.

The difficulty caused by mean shifts is that these are persistent deviations or biases from the nominal values. By persistent we mean that such a mean

Figure 8: Shifted Normal Distributions Over Tolerance Interval



shift is the same for all detail parts coming out of that process. If this mean shift has a detrimental effect on one assembly it will tend to have a similarly bad effect on all other assemblies, contingent on how it is offset by variations in the other detail parts.

The sizes and directions of these detail part mean shifts could stack in the worst possible way. A conservative approach would account for such worst case mean shifts through arithmetic stacking and stack the remaining random part to part variability via the RSS method and finally stack these two contributions (arithmetic stack of mean shifts and RSS variability stack) arithmetically. The result of this is the following tolerance stacking formula, denoting by $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)$ the set of all n mean shift ratios:

$$\begin{aligned} T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}) &= \eta_1|a_1|\,T_1 + \ldots + \eta_n|a_n|\,T_n \\ &\quad + \sqrt{(1-\eta_1)^2 a_1^2 T_1^2 + \ldots + (1-\eta_n)^2 a_n^2 T_n^2}\,. \end{aligned} \tag{6}$$

This formula is valid under independent and normal part variation with mean shifts bounded by $|\Delta_i| \leq \eta_i T_i$ and part to part variability controlled by $C_{pk} \geq 1$ for each part process.

The above combination of worst case stacking of mean shifts and RSS-stacking of the remaining variability within each tolerance interval was proposed by Mansoor (1963) and further enlarged on by Greenwood and Chase (1987).

As one can easily see, formula (6) contains our previous arithmetic and RSS stacking formulas as special cases. When $\eta_1 = \ldots = \eta_n = 0$ (no mean shift) we get

$$T_{\text{assy}}^{\Delta,\text{arith},1}(\mathbf{0}) = T_{\text{assy}}^{\text{stat}}$$

and when $\eta_1 = \ldots = \eta_n = 1$ (mean shift all the way to the tolerance limits) we get

$$T_{\text{assy}}^{\Delta,\text{arith},1}(\mathbf{1}) = T_{\text{assy}}^{\text{arith}} \, .$$

The latter occurs because in that case all deviations from nominal are represented by the mean shifts and no more part to part variation is allowed because of the $C_{pk} \geq 1$ requirement.

In general we have

$$T_{\text{assy}}^{\text{stat}} \quad \leq \quad T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}) \quad \leq \quad T_{\text{assy}}^{\text{arith}}$$

and it is worth pointing out that $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta})$ grows on the order of n, just as $T_{\text{assy}}^{\text{arith}}$, however with a reduction via the factors η_i.

Although the two main components to $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta})$, as given in (6), react differently to increases in the fraction η_i (the one increases whereas the other decreases as η_i increases) one can show (Scholz, 1995) that the overall net effect is that $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta})$ increases with η_i. Thus we can use (6) also as an upper bound for the assembly tolerance for all η_i values less than those used in (6).

From an operational point of view we can say that at least 99.73% of all assembly gaps G will fall within $\gamma \pm T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta})$. The reason for saying “at least” is that G is normally distributed and in the worst case mean shift configuration it has a mean shifted away from the nominal γ and only one tail of its distribution will significantly stick out beyond $\gamma \pm T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta})$. Figure 8 illustrates this point. By the method employed we have limited that tail probability to half of $(100 - 99.73)\%$, i.e., to $.135\%$. The opposite distribution tail will typically amount to much less than $.135\%$. How much so depends on the total amount of shift and the overall variability in G, see Scholz (1995).

In the foregoing treatment of mean shifts we had T_i fixed and traded mean shift off against part to part variability. Typically however, the $\pm 3\sigma_i$ range of the detail part process is more or less known and set and not so easily changed. Rather than taking $T_i = 3\sigma_i$ with no allowance for mean shift we would accommodate such shifts not by insisting on a σ_i reduction but by using an inflated value for T_i, namely

$$T_i = 3\sigma_i + \eta_i T_i \implies (1-\eta_i)T_i = 3\sigma_i \quad \text{or} \quad T_i = \frac{3\sigma_i}{1-\eta_i} \ ,$$

in order to absorb a mean shift of absolute size $\le \eta_i T_i$ while maintaining a $C_{pk} \ge 1$.

If we use the widened tolerances $T_i = 3\sigma_i/(1-\eta_i)$, and write $T_i' = 3\sigma_i$ for short, we can rewrite (6) as follows

$$\boxed{\begin{aligned} T_{\text{assy}}^{\Delta,\text{arith},2}(\boldsymbol{\eta}) \;\; &= \;\; \frac{\eta_1}{1-\eta_1}|a_1|\, T_1' + \ldots + \frac{\eta_n}{1-\eta_n}|a_n|\, T_n' \qquad (7) \\ &\qquad + \sqrt{a_1^2 T_1'^2 + \ldots + a_n^2 T_n'^2}\ . \end{aligned}}$$

Note that the second term is the usual RSS stack of part to part variation and the first represents the arithmetic mean shift stack expressed in terms of that part to part variability $T_i' = 3\sigma_i$. This formula is valid under independent and normal part variation with mean shifts bounded by $|\Delta_i| \le \eta_i T_i'/(1-\eta_i)$ and $T_i' = 3\sigma_i$ characterizes the part to part variability for the i^{th} part.

Here it is quite obvious that $T_{\text{assy}}^{\Delta,\text{arith},2}(\boldsymbol{\eta})$ increases with η_i since $\eta_i/(1-\eta_i)$ is increasing in η_i.

The discussion of the proportion of assemblies falling out of tolerance is completely parallel to that for $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta})$ and is thus not repeated here for $T_{\text{assy}}^{\Delta,\text{arith},2}(\boldsymbol{\eta})$.

3.5 Nonnormal RSS With Arithmetically Stacked Mean Shifts

The above method of accounting for mean shifts in conjunction with normal part to part variation can be blended with our previous treatment of centered nonnormal distributions. As pointed out before, when choosing a nonnormal distribution for part to part variation one should have a good reason for doing so. Usually such nonnormal distributions will only be invoked for a few detail parts because of the accompanying penalty factors.

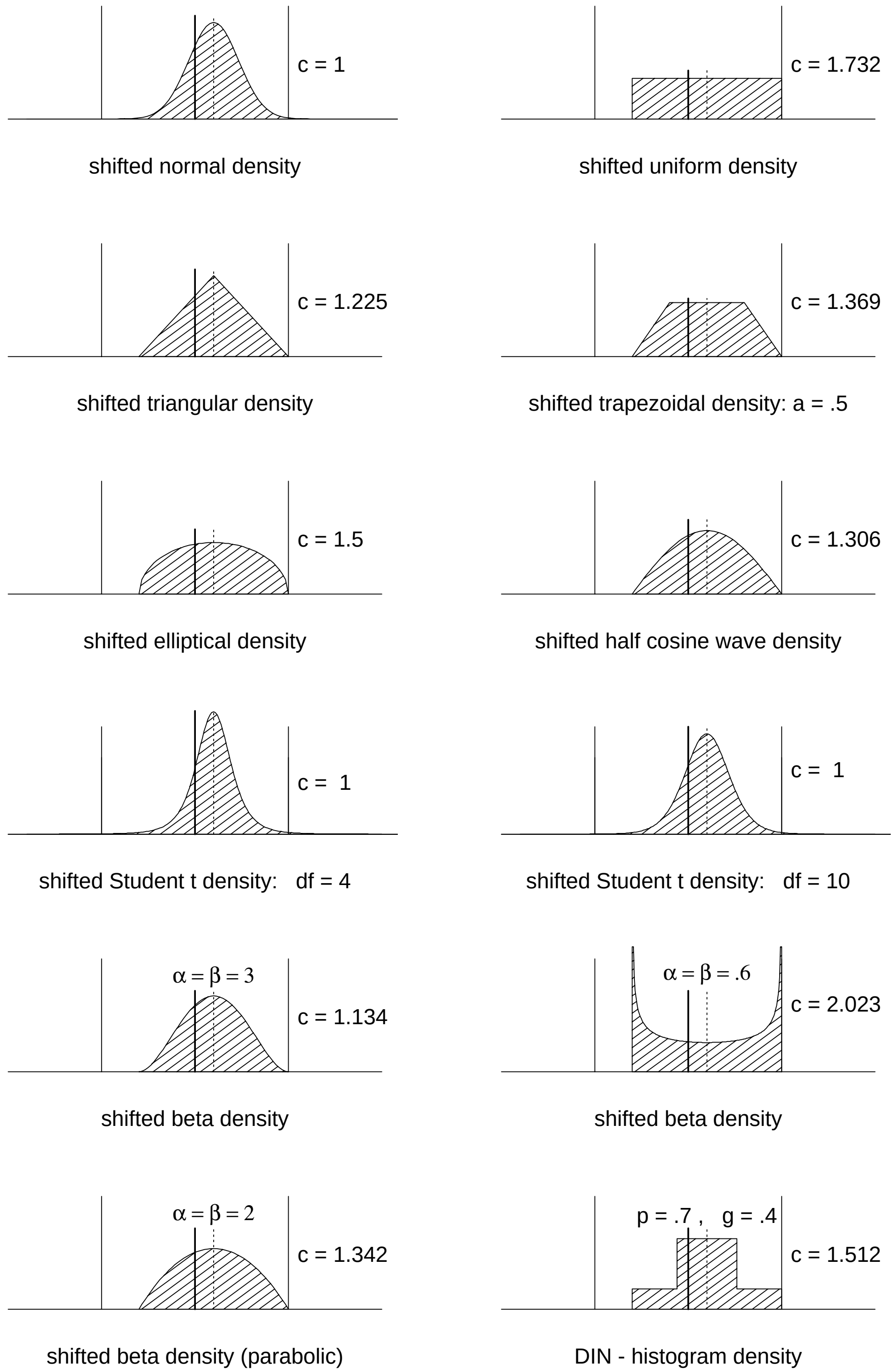
Above we allowed mean shifts as long as the total part variability including the mean shift stays mostly within the tolerance limits. This was expressed by the requirement $C_{pk} \geq 1$. The C_{pk} capability measure is strongly linked to the normal distribution. In the case of the nonnormal distributions considered previously we will thus have to reinterpret this requirement. For distributions which spread over a fixed interval we require that these distributions, after being shifted, will at most spread to the nearest endpoint of the tolerance interval. This will require that these distributions either reduce their variability around their respective means μ_i, see Figure 9 for illustrations[7], or we have to increase T_i to accommodate both Δ_i and the given fixed distributional spread of the detail part dimensions. We first state the tolerance stack formula for fixed T_i

$$
\boxed{
\begin{aligned}
T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}, \boldsymbol{c}) &= T_{\text{assy}}^{\Delta,\text{arith},1}(\eta_1, \ldots, \eta_n, c_1, \ldots, c_n) \qquad (8) \\
&= \eta_1 |a_1|\, T_1 + \ldots + \eta_n |a_n|\, T_n \\
&\qquad + \sqrt{[(1-\eta_1)c_1 a_1 T_1]^2 + \ldots + [(1-\eta_n)c_n a_n T_n]^2}
\end{aligned}
}
$$

where the factors $\boldsymbol{c} = (c_1, \ldots, c_n)$ are the same as in the centered case and can be found in Table 1. The assumption behind (8) are the same as those behind (6), except that we now allow for nonnormal part to part variation, as indicated by the choice of $\boldsymbol{c}$.

[7]note that the distributions in Figure 9 correspond to the centered distributions of Figure 6

Figure 9: Distribution Inflation Factors



As remarked above, most of the c_i will usually be one and formula (8) should be viewed as a useful extension of (6) for the occasional situation where some other than normal distribution is indicated for a detail part.

In the discussion of formula (6) we pointed out that the assembly tolerance stack, given by (6), is an increasing function of each η_i. This allowed us to treat (6) as an upper bound for all mean shift ratios which are less than those used in (6). In the case of formula (8) such claims are not possible, i.e., increasing the amount of mean shift will not necessarily increase the assembly tolerance stack $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}, \boldsymbol{c})$. The reason for this is not entirely clear and this issue could benefit from some more research. It could be that having a uniform distribution spread over $\lambda_i \pm T_i$ (i.e., $\eta_i = 0$) is more detrimental to $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}, \boldsymbol{c})$ than having a shifted uniform distribution which is more concentrated. Maybe the differences are generally small or in most situations of practical importance one still can count on $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}, \boldsymbol{c})$ being increasing in η_i so that we can again view it as upper bound for all smaller η_i. One easy check on $T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}, \boldsymbol{c})$ being increasing in η_i or not is to try out $\eta_i = 0$ and see whether that results in a larger value than obtained under $\boldsymbol{\eta}$.

Again we can claim that at least 99.73% of all assembly G values will fall within $\gamma \pm T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta}, \boldsymbol{c})$. The explanation for "at least" is as in the context of formula (6).

We now discuss the other point of view where the part variation is fixed and given, i.e., lies mostly within $\mu_i \pm T_i'$. Here $T_i' = 3\sigma_i$ (in the case of the normal or Student t-distribution) or $T_i' =$ half width of the interval over which a finite range distribution spreads out (e.g. uniform, triangular, etc.). In order to accommodate not only that variation but also the permitted amount of mean shift we will have to use a T_i that is inflated relative to T_i', namely

$$T_i = \eta_i T_i + T_i' \quad \text{or} \quad T_i = \frac{T_i'}{1 - \eta_i} \; .$$

The resulting assembly stack formula derives from (8) as

$$
\begin{aligned}
T_{\text{assy}}^{\Delta,\text{arith}}(\boldsymbol{\eta}, \boldsymbol{c}) &= T_{\text{assy}}^{\Delta,\text{arith}}(\eta_1, \ldots, \eta_n, c_1, \ldots, c_n) \qquad (9) \\
&= \frac{\eta_1}{1-\eta_1}\, |a_1|\, T_1' + \ldots + \frac{\eta_n}{1-\eta_n}\, |a_n|\, T_n' \\
&\quad + \sqrt{(c_1 a_1 T_1')^2 + \ldots + (c_n a_n T_n')^2}
\end{aligned}
$$

where the second part bears a strong resemblance to the RSS stacking formula (5) for nonnormal part to part variation and the first part is the arithmetically stacked mean shift contribution. Note again that in this formulation the assembly stack is an increasing function of the η_i.

The above stacking formula (9) is valid under the same assumptions as formula (7), except that now we allow nonnormal part variation as indicated by the c factors.

Yet another method of dealing with mean shifts in a worst case fashion is presented in Srinivasan et al. (1995). However, their results are more concerned with worst case risk for given part tolerances but still need translation into assembly tolerance stack form for given risk in order to make them comparable with the methods presented here.

3.6 RSS Stacked Part Variation and Mean Shifts I

We have seen that the simple RSS method led to a significant reduction in the assembly tolerance stack when compared to simple arithmetic stacking. The latter grows at the rate of n whereas the former grows at the much slower rate of $\sqrt{n}$. The RSS method was applicable under the assumption that the part variations were centered on nominal, random and statistically independent over the n parts in the assembly. This advantage was gained from the realization that the deviations from nominals would cancel each other out to some extent and were not likely to stack in the worst possible way.

The same considerations could lead us to realize a similar gain by stacking mean shifts via the RSS method and combine that arithmetically with the

RSS stack of the part to part variation. The assumptions are that the mean shifts are random over the specified intervals $[\lambda_i - \eta_i T_i, \lambda_i + \eta_i T_i]$ and are statistically independent from part process to part process. The randomness of each mean shift is described by some distribution, such as a normal or t-distribution with $3\sigma_\mu = \eta_i T_i$, or a uniform, triangular or any of the other finite range distributions with range width $= 2\eta_i T_i$. The choice of distribution for the mean shift is indicated by a c constant, denoted by $c_{\mu,i}$ in this context and chosen again from Table 1.

In this section we assume that the part variation is fixed and given by T_i' and that the maximum absolute mean shift is bounded by $\eta_i T_i$, with T_i yet to be determined to accommodate both part variation and maximum allowable mean shift. We get again

$$T_i = \eta_i T_i + T_i' \quad \text{or} \quad T_i = \frac{T_i'}{1-\eta_i}$$

as the inflated tolerance. However, here the actual mean shift will typically not attain its allowed maximum $\eta_i T_i$ but will vary randomly over the interval $\pm\eta_i T_i$ according to the distributions of choice as indicated by $c_{\mu,i}$ taken from Table 1. The crucial assumption is that this mean shift variation is independent from part process to part process.

Under these assumptions and denoting by $\boldsymbol{c}_\mu = (c_{\mu,1}, \ldots, c_{\mu,n})$ the inflation factors for the mean shift distributions we statistically stack these n mean shifts to bound the aggregate mean shift, as it propagates through to the assembly, by

$$\begin{aligned}&\sqrt{c_{\mu,1}^2 a_1^2 \eta_1^2 T_1^2 + \ldots + c_{\mu,n}^2 a_n^2 \eta_n^2 T_n^2} \\ &\qquad = \sqrt{c_{\mu,1}^2 a_1^2 \eta_1^2 T_1'^2/(1-\eta_1)^2 + \ldots + c_{\mu,n}^2 a_n^2 \eta_n^2 T_n'^2/(1-\eta_n)^2}\,.\end{aligned}$$

This should bound 99.73% of all assembly mean shifts. Assuming a worst case stance, namely taking the above bound as the worst assembly mean shift, and adding to that (in worst case fashion) the RSS stacked part variation

$$\begin{aligned}&\sqrt{c_1^2 a_1^2 (1-\eta_1)^2 T_1^2 + \ldots + c_n^2 a_n^2 (1-\eta_n)^2 T_n^2} \\ &\qquad = \sqrt{c_1^2 a_1^2 T_1'^2 + \ldots + c_n^2 a_n^2 (1-\eta_n)^2 T_n'^2}\end{aligned}$$

we get as assembly tolerance stack the following formula

$$
\boxed{
\begin{aligned}
&T_{\text{assy}}^{\Delta,\text{stat},1}(\boldsymbol{\eta},\boldsymbol{c},\boldsymbol{c}_\mu) \qquad\qquad (10)\\
&= \sqrt{c_{\mu,1}^2 a_1^2\eta_1^2T_1^2+\ldots+c_{\mu,n}^2a_n^2\eta_n^2T_n^2}\\
&\qquad\qquad + \sqrt{c_1^2a_1^2(1-\eta_1)^2T_1^2+\ldots+c_n^2a_n^2(1-\eta_n)^2T_n^2}\\
&= \sqrt{c_{\mu,1}^2a_1^2\eta_1^2T_1'^2/(1-\eta_1)^2+\ldots+c_{\mu,n}^2a_n^2\eta_n^2T_n'^2/(1-\eta_n)^2}\\
&\qquad\qquad + \sqrt{c_1^2a_1^2T_1'^2+\ldots+c_n^2a_n^2(1-\eta_n)^2T_n'^2}\,.
\end{aligned}
}
$$

Formula (10) is valid provided that i) the mean shifts can be viewed as independent, one-time random realizations from some distributions characterized by c constants $c_{\mu,1},\ldots,c_{\mu,n}$ and mostly contained in the intervals $\pm\eta_iT_i$, and ii) the not necessarily normal part to part variation around the realized part process means μ_i is independent from part to part and is characterized by c constants $c_1,\ldots,c_n$. Here the means μ_i deviate from the nominals λ_i by the mean shifts as controlled by i). Both $c_{\mu,i}$ and c_i can be found in Table 1 for the distributions appropriate in each case.

Typically we will mostly have $c_i=1$, but concerning $c_{\mu,i}$ the error distribution for centering the part manufacturing process may be harder to choose. The reason is that usually one can get many parts from a process and measure their characteristics to establish a distribution for part to part variability, but the number of times that such processes are set up and result in new mean shifts is usually too limited to establish some meaningful distribution. Since such sparsity of hard data is a problem one may make a subjective choice and either be somehat optimistic and assume a normal distribution for all mean shifts, i.e., $c_{\mu,i}=1$, or one may conservatively take a uniform distribution over $\pm\eta_iT_i$, i.e., $c_{\mu,i}=\sqrt{3}=1.732$ for all part processes.

The nice part of formula (10) is that the mean shifts and the part variation both contribute only on the order of $\sqrt{n}$.

We need to comment on the summing of the two square root terms in (10) as opposed to using the RSS method on all terms, i.e., summing all squares under one square root. The variation due to mean shifts is a one time affair. Once the part processes are set, with whatever random mean shifts they experienced, it is assumed that they will stay at those mean shifts.[8] The part to part variations around these set means μ_i happen anew for each part produced. The RSS part of the mean shifts bounds the total assembly mean shift probabilistically with high assurance $\approx 99.73\%$. However, we do not know where in that bounded range that assembly mean shift lies. It sits there and no longer moves. It this is difficult to give it any long run frequency interpretation within that assembly setup. To be conservative and safe we assume the worst, namely that the assembly mean shift lies at one end of the bounded range, namely at

$$\gamma \pm \sqrt{c_{\mu,1}^2 a_1^2 \eta_1^2 T_1^2 + \ldots + c_{\mu,n}^2 a_n^2 \eta_n^2 T_n^2}\ ,$$

and add to that the RSS bounded part to part variation. This step of adding is a form of worst case analysis. We do not expect variation cancellation between assembly mean shift and aggregate part variation, because that assembly mean shift does not vary from assembly to assembly whereas the aggregate part variation does vary.

Some insight into formula (10) is gained by setting $\eta_i = \eta$ (same mean shift fraction for all parts) and $c_i = c$ and $c_{\mu,i} = c_\mu$. Then

$$\begin{aligned} T_{\text{assy}}^{\Delta,\text{stat},1}(\boldsymbol{\eta}, \boldsymbol{c}, \boldsymbol{c}_\mu) &= [\eta c_\mu/(1-\eta) + c]\sqrt{a_1^2 T_1'^2 + \ldots + a_n^2 T_n'^2} \\ &= [\eta c_\mu + c(1-\eta)]\sqrt{a_1^2 T_1^2 + \ldots + a_n^2 T_n^2} \end{aligned}$$

where the RSS term in the first line is just as in the centered case (without mean shift) and the multiplier $(\eta c_\mu/(1-\eta) + c)$ adjusts not only for possible nonnormality in part and mean shift variation but also for the presence of a mean shift in itself. The latter becomes more apparent when we assume $c = c_\mu = 1$ when that multiplier becomes $(1+\eta/(1-\eta))$. It is interesting that in this simplified setup mean shifts also are compensated for by an inflation factor. The second line of (10) expresses the tolerance stack in terms of the inflated T_i, as they are used for part tolerance specification.

[8]this does not allow for part processes that have a drifting mean as the parts are produced

For later comparison we consider the following example scenario: $\eta_i = .2$, $c_i = 1$ and $c_{\mu,i} = \sqrt{3}$ for all i, i.e., common .2 mean shift ratio, common normal part variation and common uniform mean shift distributions. Then

$$[\eta c_\mu/(1-\eta) + c] = 1.433 \quad \text{or} \quad [\eta c_\mu + c(1-\eta)] = .8 \times 1.433 = 1.146 \; .$$

The assembly tolerance stack in formula (10) is of order $\sqrt{n}$, i.e., similar to the RSS method with an inflation factor, however the motivation is different in that we allow and account for some amount of mean shift. Also, the proportion of assembly gaps G falling within $\pm T_{\text{assy}}^{\Delta,\text{stat},1}(\boldsymbol{\eta}, \boldsymbol{c}, \boldsymbol{c}_\mu)$ of γ is at least .9973 and more likely at least .99865.

When speaking of at least 99.865% assurance for assembly gaps G to be within tolerance of γ, we are assuming that the actual assembly mean shift $\sum_{i=1}^n a_i \Delta_i$ is within

$$\pm\sqrt{c_{\mu,1}^2 a_1^2 \eta_1^2 T_1^2 + \ldots + c_{\mu,n}^2 a_n^2 \eta_n^2 T_n^2} \; .$$

However, this in itself is a chance event, namely it has .9973 chance of happening. The reason for not blending this probability with the previous one is that these two chances have different operational meaning. The chance concerning assembly mean shifts is not taken too often, thus we will rarely see such mean shifts out of the assumed tolerance. However, the chance of .99865 should be viewed against the fact that many assemblies will be produced with this particular setup of part processes. The fraction of out of tolerance assemblies will become noticeable in the long run.

3.7 RSS Stacked Part Variation and Mean Shifts II

In the previous section we dealt with a variation model that treated part to part variation as fixed and given, namely as T_i', and allowed a maximum absolute mean shift $\eta_i T_i$, where T_i was inflated to accommodate both requirements.

Here we take the view that T_i is given, with an allowed maximum absolute mean shift of $\eta_i T_i$ and the part to part variation can vary depending on the amount of mean shift realized for that particular part process. For example, if the part process has no mean shift, then the part variation can use up the entire $\pm T_i$ range, i.e., $T_i' = T_i$. If the mean shift is at its maximal value, then the part variation has to be greatly reduced, namely $T_i' = T_i - \eta_i T_i = (1-\eta_i)T_i$. We don't say that this dynamic behavior of part variation will happen, but if we specify the part processes through the mean shift $|\Delta_i| \le \eta_i T_i$ and $C_{pk} \ge 1$ requirements, then we do leave ourselves open to just such contingencies. The proper treatment of tolerance stacking in this case is more complicated and not quite as broadly developed as the previous methods.

Breaking mildly away from not burdening the reader with theoretical details we will here give some theoretical insight into this particular tolerance stacking method. To this end we introduce some useful notation, namely the random fraction of mean shift:

$$R_i = \frac{\Delta_i}{\eta_i T_i} = \frac{\mu_i - \lambda_i}{\eta_i T_i}$$

which is considered to vary randomly over the interval $[-1, 1]$. The random values $\boldsymbol{R} = (R_1, \ldots, R_n)$ specify the realized relative amounts of mean shift for the n parts. With this notation and recalling $\epsilon_i = L_i - \mu_i$ and $\Delta_i = \mu_i - \lambda_i$ we can write

$$G - \gamma = \sum_{i=1}^n a_i(L_i - \lambda_i) = \sum_{i=1}^n a_i \Delta_i + \sum_{i=1}^n a_i \epsilon_i = \sum_{i=1}^n a_i \eta_i T_i R_i + \sum_{i=1}^n a_i \epsilon_i \ .$$

Here the first sum on the right reflects the assembly mean shift and the second sum the assembly variation from part to part. We permit up to $\pm\eta_i T_i$ mean shift, but it may be far less depending on the mean shift reduction factors R_i. Conservatively we allow that the maximal part variability could use up as much as is made possible by the actually realized mean shift fraction. Namely, for fixed R_i the variability of the ϵ_i terms can have maximal standard deviation

$$\sigma(R_i) = c_i(T_i - |\Delta_i|)/3 = c_i(1 - |R_i|\eta_i)T_i/3.$$

Note that the maximal $\sigma(R_i)$ becomes larger as $|R_i|$ gets smaller. This means that we permit more part to part variability the more centered the part process turns out to be. We do not view this as cause and effect, but more as a conservative stance of what is permitted under the rules.

For fixed values of $\boldsymbol{R} = (R_1, \ldots, R_n)$ we can again appeal to the central limit theorem and consider $G - \gamma$ to be approximately normally distributed with mean and standard deviation given respectively by

$$\sum_{i=1}^{n} a_i \eta_i T_i R_i \qquad \text{and} \qquad \frac{1}{3}\sqrt{\sum_{i=1}^{n} a_i^2 T_i^2 c_i^2 (1-|R_i|\eta_i)^2}\ .$$

Hence for fixed $\boldsymbol{R}$ we can capture 99.73% of the $G-\gamma$ variability within

$$\sum_{i=1}^{n} a_i \eta_i T_i R_i \pm \sqrt{\sum_{i=1}^{n} a_i^2 T_i^2 c_i^2 (1-|R_i|\eta_i)^2}\ .$$

Note that this interval is not centered on γ. Its position and width depend on $\boldsymbol{R}$. To cover most contingencies we need to find out how far to the right (left) the upper (lower) endpoint of this interval could typically reach as the R_i values vary. Focussing on the upper endpoint we can write

$$\begin{aligned}
&\sum_{i=1}^{n} a_i \eta_i T_i R_i + \sqrt{\sum_{i=1}^{n} a_i^2 T_i^2 c_i^2 (1-|R_i|\eta_i)^2} \\
&\qquad = \left(\sum_{i=1}^{n} w_i \eta_i R_i + \sqrt{\sum_{i=1}^{n} w_i^2 c_i^2 (1-|R_i|\eta_i)^2}\right)\sqrt{\sum_{i=1}^{n} a_i^2 T_i^2} \\
&\qquad = F(\boldsymbol{R})\sqrt{\sum_{i=1}^{n} a_i^2 T_i^2}\ ,
\end{aligned}$$

where

$$F(\boldsymbol{R}) = \sum_{i=1}^{n} w_i \eta_i R_i + \sqrt{\sum_{i=1}^{n} w_i^2 c_i^2 (1-|R_i|\eta_i)^2} \qquad \text{and} \qquad w_i = \frac{a_i T_i}{\sqrt{\sum_{i=1}^{n} a_j^2 T_j^2}}$$

Note that this is just the ordinary simple RSS formula multiplied by some allowance factor $F(\boldsymbol{R})$. Unfortunately this latter factor depends on the unknown $\boldsymbol{R}$ and it is not easy to bound this factor in simple form for all contingencies. However, it is a very simple matter to write a small simulation program that establishes such bounds for any assumptions one cares to make about η_i, c_i, T_i, a_i, and the distributions governing R_i.

For a special but also central case we can give explicit bounds for this factor. Assuming that all R_i are independent random variables, uniformly[9] distributed over $[-1,1]$ and assuming[10] $c_i = 1$ for all i and common mean shift fraction bounds $\eta_1 = \ldots = \eta_n = \eta$, we can bound $F(\boldsymbol{R})$ with high probability, namely

$$P\left(F(\boldsymbol{R}) \le \sqrt{1-\eta+\eta^2/3} + \eta\sqrt{3}\right) = .99865 \,,$$

resulting in the following tolerance stacking formula

$$\boxed{\begin{aligned} T_{\text{assy}}^{\Delta,\text{stat},2}(\eta) &= \left(\sqrt{1-\eta+\eta^2/3} + \eta\sqrt{3}\right) \\ &\qquad \times \sqrt{a_1^2T_1^2 + \ldots + a_n^2T_n^2}\,. \end{aligned}} \tag{11}$$

For $\eta = .2$ formula (11) reduces to

$$T_{\text{assy}}^{\Delta,\text{stat},2}(.2) = 1.248\,\sqrt{a_1^2T_1^2 + \ldots + a_n^2T_n^2}\,.$$

Here the factor 1.248 is somewhat larger than the factor of 1.146 derived in the previous section under the same distributional assumption, namely uniform mean shift distribution and normal part to part variation and common .2 mean shift fraction.

If we make the additional assumption that the a_iT_i are all the same, say equal to T, then we can compare the above also against the tolerance stacking method (8) which stacks the mean shifts in worst case fashion.

$$\begin{aligned} T_{\text{assy}}^{\Delta,\text{arith},1}(\boldsymbol{\eta},\mathbf{1}) &= .2\,n\,T + .8\,T\,\sqrt{n} \\ &= (.2\sqrt{n}+.8)\,T\sqrt{n} \\ &= (.2\sqrt{n}+.8)\,\sqrt{\sum_{i=1}^n a_i^2T_i^2}\,. \end{aligned}$$

[9]this is somewhat conservative, as outlined on previous occasions

[10]implied by a normal distribution for part to part variation

Thus the comparable factor of interest is $(.2\sqrt{n} + .8)$ which is 1.083, 1.146, 1.200, 1.247, ... for $n = 2, 3, 4, 5, \ldots$. By comparison these still look quite favorable, but note that $(.2\sqrt{n}+.8)$ grows like $\sqrt{n}$ without bounds, although tempered by the mean shift fraction .2, whereas the factors in the other two examples do not grow with n.

The above numbers should throw some light on the relative merits of these methods. It should also be clear by now that it matters what assumptions one makes concerning the various sources of variation. This still does not make it easy to choose and one possible compromise is to calculate tolerance stacks by two methods, say (8) and (10) or (11), and average the two results.

3.8 Risk Analysis

In all the above we have presented tolerance stacking formulas from various points of view. We have not dwelled much on the associated risks except for a few remarks made here and there. Typically the aim was to maintain the traditional risk of .27% out of compliance assemblies. This was usually related to a $\pm 3\sigma$ normal distribution range based on a central limit theorem approximation to the assembly gap. As mean shifts came into play there was some gain in realizing that the excess over assembly tolerance bounds would usually occur only at one end of the tolerance range, typically with half the risk, namely .135%. One can take advantage of this risk reduction by reducing the appropriate part of the assembly tolerance. Usually this means that the RSS part of the part to part variation should be multiplied by the factor .927. To illustrate this consider the tolerance stacking formula (8) where we place the factor .927 in front of the part variation RSS term to obtain

$$
\begin{aligned}
T_{\text{assy}}^{\Delta,\text{arith,r}}(\boldsymbol{\eta}, \boldsymbol{c}) &= T_{\text{assy,r}}^{\Delta,\text{arith}}(\eta_1, \ldots, \eta_n, c_1, \ldots, c_n) \qquad (12) \\
&= \eta_1 |a_1|\, T_1 + \ldots + \eta_n |a_n|\, T_n \\
&\quad + .927\, \sqrt{[(1-\eta_1)c_1 a_1 T_1]^2 + \ldots + [(1-\eta_n)c_n a_n T_n]^2}\,.
\end{aligned}
$$

The factor .927 is motivated by $2.782/3 = .927$, the fact that .27% of the normal curve exceeds the value 2.782, and that the RSS term originally stood for a 3σ value. For more details we refer the reader to Scholz (1995) where some of these improvements are discussed.

References

Altschul, R.E. and Scholz, F.W. (1994). "Case study in statistical tolerancing." *Manufacturing Review of the AMSE* **7**, 52-56.

ASME Y14.5M-1994, Dimensioning and Tolerancing, The American Society of Mechanical Engineers.

ASME Y14.5.1M-1994, Mathematical Definition of Dimensioning and Tolerancing Principles, The American Society of Mechanical Engineers.

Bates, E.L. (1947). "How to increase tolerances and obtain closer fits." *The Iron Age*, July 3rd.

Bates, E.L. (1949). "Specifying design tolerances." *Machine Design*, March.

Bender, A. Jr. (1962). "$6 \times 2.5 = 9$, Benderizing tolerances–a simple practical probability method of handling tolerances for limit-stack-ups." *Graphic Science*, 17-21.

Bjørke, Ø, (1989). *Computer-Aided Tolerancing*, ASME Press, New York.

Cox, N.D. (1986). *Volume 11: How to Perform Statistical Tolerance Analysis.* American Society for Quality Control, 230 West Wells Street, Milwaukee, Wisconsin 53203.

Epstein, B. (1946). "Tolerances on assemblies." *The Machinist*, April 20th.

Ettinger & Bartky (1936). "Basis for determining manufacturing tolerances." *The Machinist*, October 3rd.

Evans, D.H. (1975). "Statistical tolerancing: the state of the art. Part III. shifts and drifts." *Journal of Quality Technology* **7**, 72-76.

Fortini, E.T. (1967). *Dimensioning for Interchangeable Manufacture.* Industrial Press Inc., New York, N.Y.

Gilson, J. (1951). *A New Approach to Engineering Tolerances*, The Machinery Publishing C. LTD, 83-117 Euston Road, London, NW1.

Gladman, C.A. (1945). "Drawing office practice in relation to interchangeable components." *Proc. I. Mech. E.* **152**, No. 4, p. 388, paper and discussion.

Greenwood, W.H. and Chase, K.W. (1987). "A new tolerance analysis method for designers and manufacturers." *Trans. ASME, J. of Engineering for Industry* **109**, 112-116.

Gramenz, K. (1925). "Die Dinpassungen und ihre Anwendungen." Dinbuch 4.

Griess, K.H. (1990). "Tolerancing-Design Guide." (proprietary) The Boeing Company, D6-25382-90.

Harry, M.J. and Stewart, R. (1988). *Six Sigma Mechanical Design Tolerancing*. Motorola Government Electronics Group, 8201 E. McDowell Rd., Scottdale, AZ 85257, Ph. (602) 990-5716.

Henzold, G. (1995). *Handbook of Geometrical Tolerancing, Design, Manufacturing and Inspection*, John Wiley & Sons, New York, N.Y.

Kirschling, G. (1988). *Qualitätssicherung und Toleranzen*, Springer-Verlag, Berlin.

Loxham, J. (1947). "An experiment in the use of a standard limit system." *Proc. I. Mech. E.* **156**, No. 2, p. 103, paper and discussion.

Mansoor, E.M. (1963). "The application of probability to tolerances used in engineering designs." *Proc. of the Institution of Mechanical Engineers* **178**, 29-51 (with discussion).

Nielson, L.M. (1948). "Shop run tolerances." *Product Engineering*, May.

Nigam, S.D. and Turner, J.U. (1995). "Review of statistical approaches to tolerance analysis." *Computer-Aided Design* **27**, 6-15.

Rice, W.B. (1944). "Setting tolerances scientifically." *Mechanical Engineering*, December.

Scholz, F.W. (1995). "Tolerance stack analysis methods, a critical review." *ISSTECH-95-021* Boeing Information & Support Services.

Srinivasan, V., O'Connor, M.A., and Scholz, F.W. (1995). "Techniques for composing a class of statistical tolerance zones." *ISSTECH-95-022* Boeing Information & Support Services.

Wade, O.R. (1967). *Tolerance Control in Design and Manufacturing*, Industrial Press Inc., 200 Madison Avenue, New York 10016