https://doi.org/10.51483/IJPAMR.3.1.2023.11-45 ISSN: 2789-9160

SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Pure and Applied Mathematics Research

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Shape Factor Asymptotic Analysis II

Frank Xuyan Wang[1*]

[1]Validus Research Inc., 187 King Street South Unit 201, Waterloo, Ontario, N2J 1R1, Canada. E-mail: frank.wang@validusresearch.com

Article Info

Volume 3, Issue 1, April 2023
Received : 11 April 2022
Accepted : 27 January 2023
Published : 05 April 2023
doi: 10.51483/IJPAMR.3.1.2023.11-45

Abstract

Probability distributions with identical shape factor asymptotic limit formulas are defined as asymptotic equivalent distributions. The GB1, GB2, and Generalized Gamma distributions are examples of asymptotic equivalent distributions, which have similar fitting capabilities to data distribution with comparable parameters values. These example families are also asymptotic equivalent to Kumaraswamy, Weibull, Beta, ExpGamma, Normal, and LogNormal distributions at various parameters boundaries. The asymptotic analysis that motivated the asymptotic equivalent distributions definition is further generalized to contour analysis, with contours not necessarily parallel to the axis. Detailed contour analysis is conducted for GB1 and GB2 distributions for various contours of interest. Methods combing induction and symbolic deduction are crafted to resolve the dilemma over conflicting symbolic asymptotic limit results. From contour analysis build on graphical and analytical reasoning, we find that the upper bound of the GB2 distribution family, having the maximum shape factor for given skewness, is attained by the Double Pareto distribution.

Keywords: *Shape factor; Skewness; Kurtosis; Asymptotic equivalent distributions; GB1 distribution; Exp gamma distribution; Log normal distribution; GB2 distribution; Double pareto distribution; Contour analysis; Computer algebra system; Symbolic analysis*

1. Introduction

This is the continuation of the study in Wang (2020) for shape factor, which is defined in Wang (2019a) with a plan of addressing probability distribution fitting validation and explanation problem by examining shape factor global and conditional minimum or maximum formulas or curves, to other distribution families. The problems include, to name a few, why it is not able to fit a data distribution by some distribution families, and why different distribution families can fit a data distribution almost as good as each other while their PDF or even distribution domains are very different. The resolution is the asymptotic equivalent distribution, a concept arrived at while studying shape factor asymptotic limit, which delineate that distribution PDF shapes or ultimate tendencies, as well as relevant distribution parameters, are determined uniquely per each asymptotic equivalent distribution class.

* *Corresponding author: Frank Xuyan Wang, Validus Research Inc., 187 King Street South Unit 201, Waterloo, Ontario, N2J 1R1, Canada. E-mail: frank.wang@validusresearch.com*

The same symbolic and graphical techniques, as used in Wang (2020), will be applied to GB1 in this paper, THT distribution (defined in Wang, 2019b), GH distribution and other families later, for getting the boundary forms of the shape factor. The global minimum value of the shape factor will serve as the limit or bound of the conditional minimum value of the shape factor for some given parameters while keep other parameters unchanged. The limit form will act as the target or guide, and at many times possess analytical explicit and simpler formulas. A more delicate combined use of the induction and symbolic analysis, algebraic manipulation strategy of division and divide as well as transform, enable us to resolve a dilemma left of GB2 in Wang (2020) section 6.3.

We will refer freely on contents of Wang (2019a and 2020) rather than repeat it here, to avoid self-copying. Distributions definitions and parameter assumptions will refer to Marichev and Trott (2013). Background on GB1 and GB2 can be found in McDonald (1984), McDonald *et al.* (2011), GH and its sub-families such as VG and NIG can refer to the citations in Wang (2018b), THT and the like distributions in Wang (2019b).

Before we start our immense study of GB1 and GB2 shape factor, we will correct an error in Wang (2020) section 3.1 about limit of shape factor along skewness contour for Kumaraswamy distribution. We need to exchange the words maximum and minimum in paragraph two of its section 3.1, i.e., the conclusion should be the minimum shape factor is at the top boundary and the maximum shape factor is at the left boundary. In its section 3.1 paragraph three and Figure 3 the word maximum should be read as minimum.

Probability distribution is the core for stochastic modeling. What type of distribution is more suitable for an application is one of the fundamental questions to ask. Kurtosis, as a high order characteristic of distribution, bring about some confusions: people are indecisive about whether it measures more of the peakedness of the distribution near its center or more of the tail concentration away from distribution center. Some researchers even conclude that the kurtosis does not provide any useful information other than the mean and standard deviation and is practically worthless. However, from the perspective of application, usefulness is as important as or is more important than meaningfulness.

The shape factor, defined as the kurtosis divided by squared skewness, which we proposed and studied in a series of papers, can be think of as a modified or normalized kurtosis. Its range or extreme values are found to be useful for differentiate distribution families, and helpful for discovering relationship among distributions.

As a simple example for the usefulness of the shape factor, the Beta distribution, which is prevalent in reinsurance loss modeling, can be roughly described as having shape factor values in the interval [1, 1.5]. Then any empirical distribution with shape factor outside or near the endpoint of this interval is not appropriate for modeling as Beta distribution.

Our methodology here, the "extrema analysis"-find the range or extremum-can also be used to resolve the controversial about kurtosis. From numerical experiments with distributions having symmetric finite range piecewise linear probability distribution functions (PDF), we know that with central peakedness only, the upper bound of the kurtosis is 2.4. On the other hand, with only tail concentration, the upper bound of kurtosis is 1.8. However, with central peakedness and tail concentration together, the upper bound of kurtosis is infinite. So both of them and their combination contributed to the values of kurtosis, which measures the PDF overall or aggregated curvatures or concavities.

With this background on our shape factor concept, this paper posted on https://mpra.ub.uni-muenchen.de/110827, specifically contributed the induction and deduction combined technique to overcome symbolic limit errors for the shape factor study of GB1 and GB2. Another concept of our research output need to mention is the asymptotic equivalent distributions, which change the school of thought for distribution fitting from distribution family selection to focus on asymptotic equivalent distribution class, diminishing the importance of individual distribution forms. Its Figures 26 and 27 may be of general interest for one who want to find or confirm that whether some distribution families are most suitable for a specific application empirical distribution.

2. Results

2.1. GB1 Distribution Asymptotic Analysis

2.1.1. Shape Factor Boundary and Conditional Minimum Values When $\beta \to \infty$ for Fixed Parameter α or γ

As a generalization of the Kumaraswamy distribution, the GB1 distribution *GeneralizedBetaDistributionI*[α, β, γ, 1], is in "duality" with the GB2 distributions *BetaPrimeDistribution*[p, q, α, 1] with the parameters correspondence (α, β, γ)

$\leftrightarrow (p, q, \alpha)$, or in our terminology, they are asymptotically equivalent distributions, a typical scenario of a pair of very different distributions with identical shape factor. We will see soon that their boundary shape factor forms are the same and the curve of the location of shape factor conditional minimum in the parameters space are coinciding.

The trick in Equation (12) of Wang (2020) can be used to find whether an asymptotic limit of power function exist, and if exist, what the power order is.

We used the following formula in Wang (2020) section 6.3 for GB2, and it will be used for GB1 as well:

$$\lim_{p\to\infty}\left(p+\frac{z}{\alpha}\right)^{-o+n\left(p+\frac{z}{\alpha}\right)} \sim e^{\frac{nz}{\alpha}} p^{-o+n\left(p+\frac{z}{\alpha}\right)} \quad ...(1)$$

$$\lim_{\alpha\to 0}\left(p+\frac{z}{\alpha}\right)^{-o+n\left(p+\frac{z}{\alpha}\right)} \sim e^{pn}\left(\frac{z}{\alpha}\right)^{-o+n\left(p+\frac{z}{\alpha}\right)} \quad ..(2)$$

First, we have the central moment formula for GB1 when $\beta\to\infty$:

$$\lim_{\beta\to\infty} CM[2] \sim \frac{(\alpha+\beta)^{-2/\gamma}\left(-Gamma\left[\alpha+\frac{1}{\gamma}\right]^2+Gamma[\alpha]Gamma\left[\alpha+\frac{2}{\gamma}\right]\right)}{Gamma[\alpha]^2} \quad ...(3)$$

$$\lim_{\beta\to\infty} CM[3] \sim \frac{(\alpha+\beta)^{-3/\gamma}\left(2Gamma\left[\alpha+\frac{1}{\gamma}\right]^3-3Gamma[\alpha]Gamma\left[\alpha+\frac{1}{\gamma}\right]Gamma\left[\alpha+\frac{2}{\gamma}\right]+Gamma[\alpha]^2 Gamma\left[\alpha+\frac{3}{\gamma}\right]\right)}{Gamma[\alpha]^3} \quad ...(4)$$

$$\lim_{\beta\to\infty} CM[4] \sim \frac{(\alpha+\beta)^{-4/\gamma}}{Gamma[\alpha]^4}\left(-3Gamma\left[\alpha+\frac{1}{\gamma}\right]^4+6Gamma[\alpha]Gamma\left[\alpha+\frac{1}{\gamma}\right]^2 Gamma\left[\alpha+\frac{2}{\gamma}\right]\right.$$
$$\left.-4Gamma[\alpha]^2 Gamma\left[\alpha+\frac{1}{\gamma}\right]Gamma\left[\alpha+\frac{3}{\gamma}\right]+Gamma[\alpha]^3 Gamma\left[\alpha+\frac{4}{\gamma}\right]\right) \quad ...(5)$$

From them we get the boundary values:

$$\lim_{\alpha\to 0}\lim_{\beta\to\infty} SF = \frac{Gamma\left[\frac{2}{\gamma}\right]Gamma\left[\frac{4}{\gamma}\right]}{Gamma\left[\frac{3}{\gamma}\right]^2} \quad ...(6)$$

$$\lim_{\gamma\to\infty}\lim_{\beta\to\infty} SF = \frac{PolyGamma[1,\alpha]\left(3PolyGamma[1,\alpha]^2+PolyGamma[3,\alpha]\right)}{PolyGamma[2,\alpha]^2} \quad ...(7)$$

$$\lim_{\gamma\to 0}\lim_{\beta\to\infty} SF \sim \left(\frac{8}{9}\right)^{\alpha-\frac{1}{2}}\left(\frac{1024}{729}\right)^{\frac{1}{\gamma}} \quad ...(8)$$

$$\lim_{\alpha\to\infty}\lim_{\beta\to\infty} SF \sim \frac{3\alpha\gamma^2}{(-3+\gamma)^2} \quad ...(9)$$

To get Equation (9) we divide the shape factor into three parts, for each of them use derivative of their log function to find the power order, and then confirm the power function asymptotic expression by taking symbolic limit of their quotient, for example,

$$Limit\left[\frac{\frac{2Gamma\left[\alpha+\frac{1}{\gamma}\right]^3-3Gamma[\alpha]Gamma\left[\alpha+\frac{1}{\gamma}\right]Gamma\left[\alpha+\frac{2}{\gamma}\right]+Gamma[\alpha]^2Gamma\left[\alpha+\frac{3}{\gamma}\right]}{Gamma[\alpha]^2Gamma\left[\alpha+\frac{3}{\gamma}\right]}}{\frac{3-\gamma}{\alpha^2\gamma^4}},\alpha\to\infty,Direction\to\right.$$

"FromBelow", Assumptions $\to \alpha > 0 \&\& \gamma > 0] = 1$; and use Equation (1) to simplify the gamma function expansion:

$$\left.Limit\left[\frac{\frac{Gamma\left[\alpha+\frac{2}{\gamma}\right]Gamma\left[\alpha+\frac{4}{\gamma}\right]}{Gamma\left[\alpha+\frac{3}{\gamma}\right]^2}}{\frac{\left(\alpha+\frac{2}{\gamma}\right)^{\alpha+\frac{2}{\gamma}-\frac{1}{2}}\left(\alpha+\frac{4}{\gamma}\right)^{\alpha+\frac{4}{\gamma}-\frac{1}{2}}}{\left(\alpha+\frac{3}{\gamma}\right)^{2\alpha+\frac{6}{\gamma}-1}}},\alpha\to\infty,\text{Direction}\to\text{"FromBelow"},\text{Assumptions}\to\alpha>0\&\&\gamma>0\right]=1\right.$$

.

$$\left.Limit\left[\frac{\frac{e^{\frac{2}{\gamma}}(\alpha)^{\alpha+\frac{2}{\gamma}-\frac{1}{2}}e^{\frac{4}{\gamma}}(\alpha)^{\alpha+\frac{4}{\gamma}-\frac{1}{2}}}{e^{\frac{6}{\gamma}}(\alpha)^{2\alpha+\frac{6}{\gamma}-1}}}{\frac{\left(\alpha+\frac{2}{\gamma}\right)^{\alpha+\frac{2}{\gamma}-\frac{1}{2}}\left(\alpha+\frac{4}{\gamma}\right)^{\alpha+\frac{4}{\gamma}-\frac{1}{2}}}{\left(\alpha+\frac{3}{\gamma}\right)^{2\alpha+\frac{6}{\gamma}-1}}},\alpha\to\infty,\text{Direction}\to\text{"FromBelow"},\text{Assumptions}\to\alpha>0\&\&\gamma>0\right]=1\right.$$

.

Here another trick of dividing by the largest term or highest asymptotic order term in a sum expression instead of working on the sum expression directly is used, which enable us to overcome the indecisive results if we use serial expansion (Wang (2020) section 6.3).

With the substitution of parameters $(\alpha, \beta, \gamma) \leftrightarrow (p, q, \alpha)$, from Equation (9) we also get the boundary value of GB2 distribution:

$$\lim_{p\to\infty}\lim_{q\to\infty} GB2SF \sim \frac{3p\alpha^2}{(-3+\alpha)^2} \qquad ...(10)$$

Equation (10) solved the puzzle we left in Wang (2020): it says the limit is positive infinite rather than the asymptotic expansion results of $\lim_{p\to\infty} SFB \sim \frac{4}{3}$. It cautions about the use of asymptotic series expansions or unsuspicious confidence in symbolic limit (see also https://web.archive.org/web/20190425210704/https://www.linkedin.com/pulse/pitfall-symbolic-limit-wang-frank/). It also cautions about innocent trust in numerical results, for example the calculated zero values or negative values of the shape factor for huge p in packages such as MPMATH (many of the wrong results like this can be easily spotted out from the simple property that $K>1$, $SF>1$, and $K>S$).

It merits to emphasize the possible error generated through simplification using asymptotic equivalent expressions to substitute terms in a sum expression, for instance by using Equation (1) or (2). The reason for the error is that when the largest order terms in the expansion cancel out each other, the less significant terms that have been ignored become

the dominant terms, which then lead to contradictory results through different substitutions or algebraic manipulations. Thus show the 4/3 puzzle. Therefore, the symbolic analysis should not be used alone to derive formulas. It should be combined with numerical experiment to correct errors of each other. The symbolic analysis can be used as exploration

Figure 1: GB1 Skewness 3D Contour Plot at $\beta = \infty$

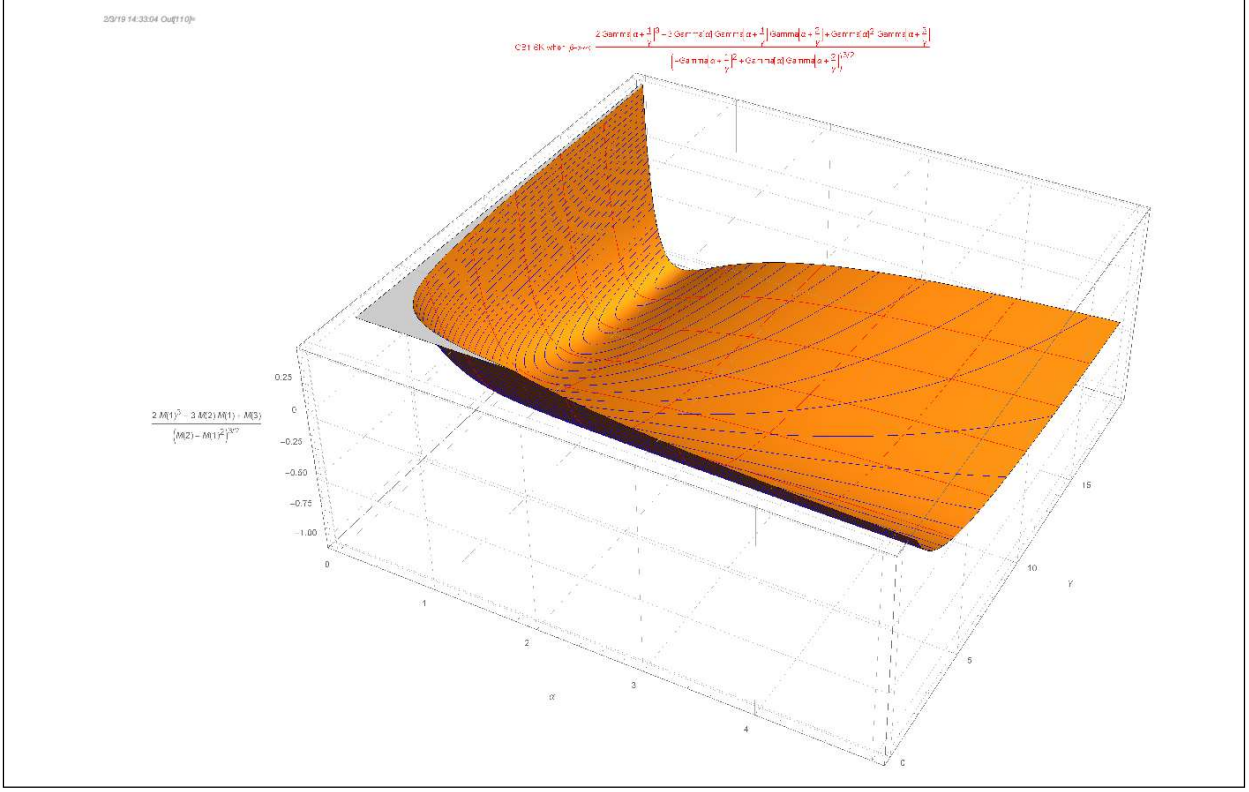
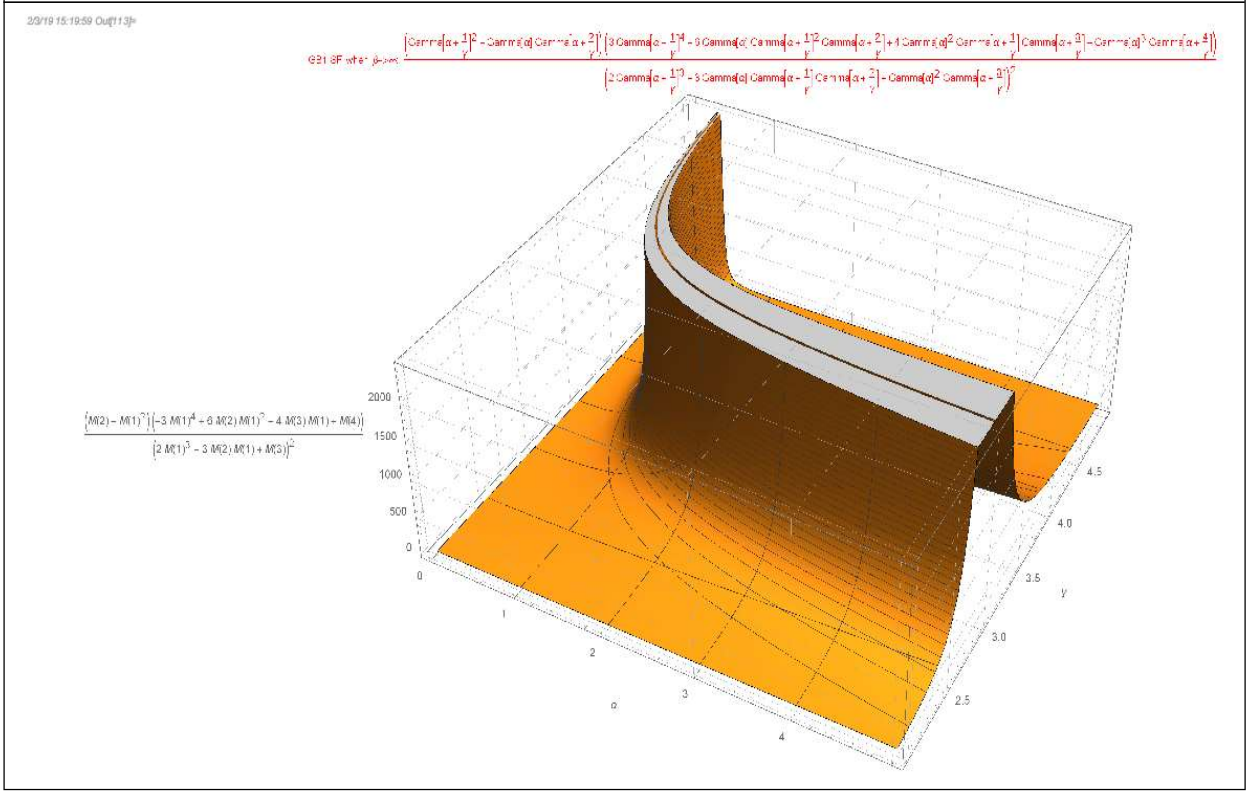


Figure 2: GB1 Shape Factor 3D Contour Plot at $\beta = \infty$



tool to provide candidate formulas for empirical study, or as verification of the empirical formulas. In case the exploration gives absurd results (such as negative or less than one *K* and *SF*), contradictory or uncertain results, there is a remedy of calculating the formula for specific parameter values without using any of such asymptotic equivalent substitutions. Generally, if deduction get us into quagmire, we can try induction instead. In summary, with no simplification, we get no results, but with simplification, we may get wrong results. Nevertheless, we will still use symbolic deductions extensively, only be alert of the possibility of overturn of the symbolic results if they have not yet been confirmed through diverse routes.

Since the uncertainty in symbolic analysis, similar to the uncertainty in numerical analysis (Wang (2019b) section 2.3 and 2.5), gives us much dilemmas, we will expound on this problem and the induction solution further by example in later section 2.2.1.

We have similar to GB2 formulas for skewness, kurtosis, and shape factor of GB1 distribution when $\beta \to \infty$:

$$\lim_{\beta\to\infty} S \sim \frac{2Gamma\left[\alpha+\frac{1}{\gamma}\right]^3 - 3Gamma[\alpha]Gamma\left[\alpha+\frac{1}{\gamma}\right]Gamma\left[\alpha+\frac{2}{\gamma}\right] + Gamma[\alpha]^2 Gamma\left[\alpha+\frac{3}{\gamma}\right]}{\left(-Gamma\left[\alpha+\frac{1}{\gamma}\right]^2 + Gamma[\alpha]Gamma\left[\alpha+\frac{2}{\gamma}\right]\right)^{3/2}} \quad ...(11)$$

$$\lim_{\beta\to\infty} K \sim \frac{1}{\left(Gamma\left[\alpha+\frac{1}{\gamma}\right]^2 - Gamma[\alpha]Gamma\left[\alpha+\frac{2}{\gamma}\right]\right)^2}\left(-3Gamma\left[\alpha+\frac{1}{\gamma}\right]^4 + 6Gamma[\alpha]Gamma\left[\alpha+\frac{1}{\gamma}\right]^2 Gamma\left[\alpha+\frac{2}{\gamma}\right]\right.$$

$$\left. -4Gamma[\alpha]^2 Gamma\left[\alpha+\frac{1}{\gamma}\right]Gamma\left[\alpha+\frac{3}{\gamma}\right] + Gamma[\alpha]^3 Gamma\left[\alpha+\frac{4}{\gamma}\right]\right) \quad ...(12)$$

$$\lim_{\beta\to\infty} SF \sim \frac{Gamma\left[\alpha+\frac{1}{\gamma}\right]^2 - Gamma[\alpha]Gamma\left[\alpha+\frac{2}{\gamma}\right]}{\left(2Gamma\left[\alpha+\frac{1}{\gamma}\right]^3 - 3Gamma[\alpha]Gamma\left[\alpha+\frac{1}{\gamma}\right]Gamma\left[\alpha+\frac{2}{\gamma}\right] + Gamma[\alpha]^2 Gamma\left[\alpha+\frac{3}{\gamma}\right]\right)^2}\left(3Gamma\left[\alpha+\frac{1}{\gamma}\right]^4\right.$$

$$\left. -6Gamma[\alpha]Gamma\left[\alpha+\frac{1}{\gamma}\right]^2 Gamma\left[\alpha+\frac{2}{\gamma}\right] + 4Gamma[\alpha]^2 Gamma\left[\alpha+\frac{1}{\gamma}\right]Gamma\left[\alpha+\frac{3}{\gamma}\right] - Gamma[\alpha]^3 Gamma\left[\alpha+\frac{4}{\gamma}\right]\right) \quad ...(13)$$

The 3D plot of the Equation (11) and (13) is in Figures 1-2.

The conditional minimum shape factor plot for given α or for given γ is in Figures 3-4.

These plots of GB1 are identical to the corresponding plots of GB2. Since Figures 3-4 will give relative good parameter estimation when matching given shape factor, shifted GB1 and GB2 distribution will have approximately identical parameters for their best fit to given data distribution. Fitting practice confirmed that the best fitted shifted GB1 and GB2 distributions *GeneralizedBetaDistributionI*[α, β, γ, b] and *BetaPrimeDistribution*[p, q, α, β] shape parameters under the correspondence (α, γ) ↔ (p, α) are very close, with the differences mainly on the fourth or scale parameters b and β respectively, which are shape factor independent. This appears to contradict the results that the GB2 shape factor have higher range than GB1 as shown by Figure 1 in McDonald *et al.* (2011). What it exactly saying is that the conditional minimum of the shape factor along different contour, or different directional minimum of the shape factor, behaves differently, similar to the characteristic number of GB2 that we discussed in Wang (2020) (the characteristic number conclusion is also valid for GB1 through the parameters transformation (α, γ) ↔ (p, α)). We will confirm this in later subsections 2.3.1, 2.4.2, and 2.5.1, by the analysis of the conditional minimum of the shape factor for given skewness, a different type of the contour.

Even though GB1 and GB2 have almost identical fitting capability and fitted parameters as we mentioned, experience indicates that GB1 is easier to use with fewer technical problems in numerical optimization performed by the fitting, and is easier for validation using risk measures that involves numerical integration (Wang, 2019b). This relatively easier assertion is generally true for finite range distributions vs. infinite range distributions.

Figure 3: GB1 Minimum SF Plot for Given α, Get from Asymptotic Expression When $\beta \to \infty$

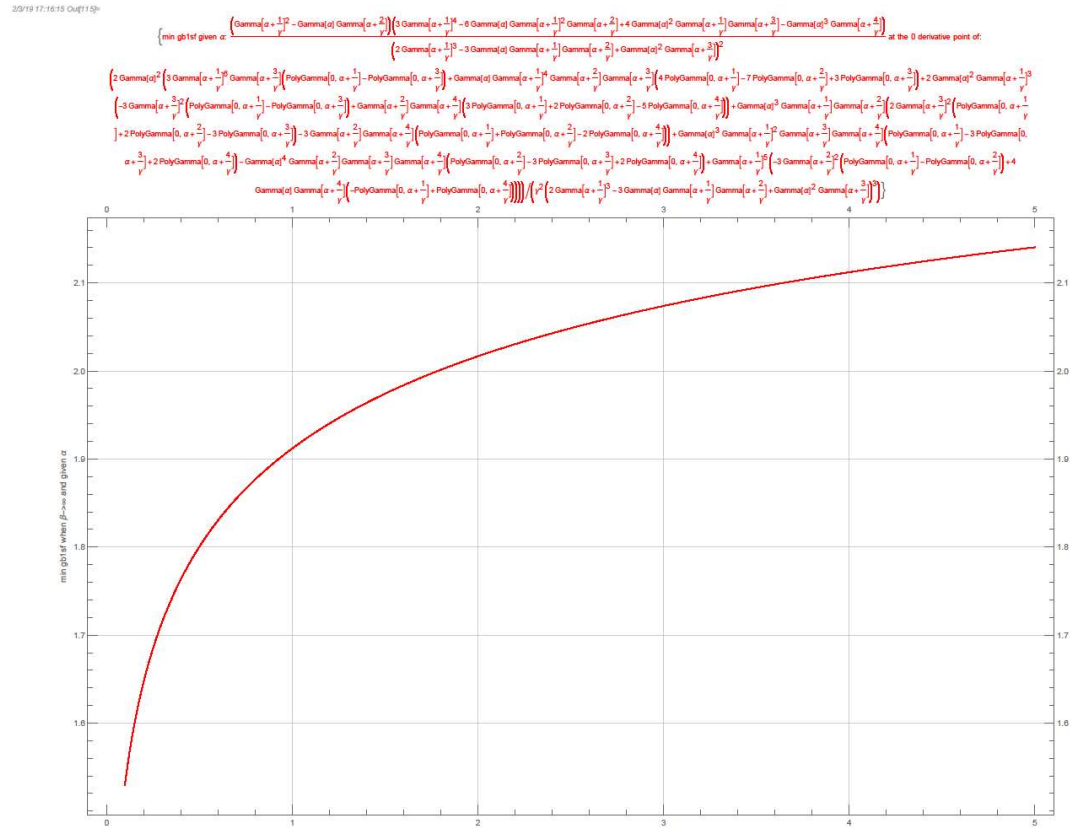
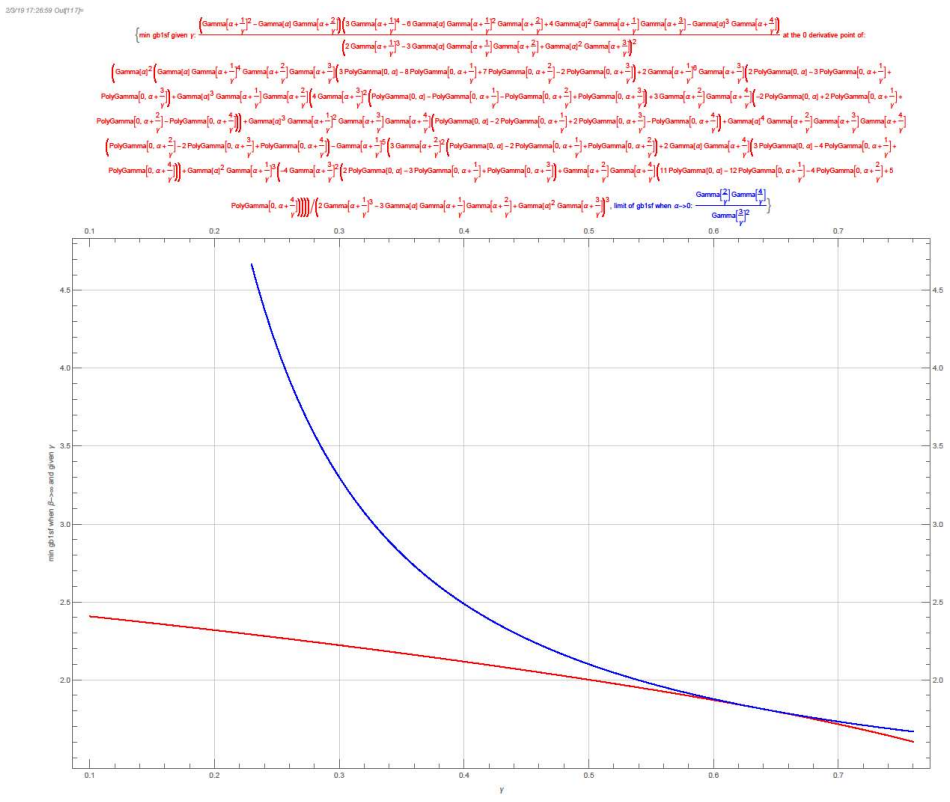


Figure 4: GB1 Minimum SF Plot When $\beta \to \infty$ for Given γ, Get from SF Zero Derivative Point Value (in Red) or Boundary Value When $\alpha \to 0$ (in Blue)



2.1.2. Asymptotic Equivalent Distributions of GB1

In Wang (2020) section 5.2, we call distributions with identical SF boundary value formulas asymptotically equivalent distributions. In this sense we see that GB1, GB2, and Generalized Gamma Distribution (GG) are asymptotic equivalent, denoting by $GeneralizedBetaDistributionI[\alpha, \infty, \gamma, 1] \sim BetaPrimeDistribution[\alpha, \infty, \gamma, 1] \sim GammaDistribution[\alpha, 1, \gamma, 0]$. Their PDF do not converge to each other, or even do not have the same domain. Their central moments do not have the same formula either, but their *S*, *K*, and *SF* have the same asymptotic formula when $\beta \to \infty$ or $q \to \infty$. Asymptotic equivalent distributions will have similar fitting capabilities and fitted parameters, in justification of our observation that in practice when Beta distribution do not fit well all the distributions in these families will not fit well. In the asymptotic equivalent notation, when $\beta \to \infty$, in the parameters space of α and γ, Beta distribution is the GB1 distribution at $\gamma = 1$: $GeneralizedBetaDistributionI[\alpha, \infty, 1, 1] \sim BetaDistribution[\alpha, \infty]$. We can similarly summarize our finding in Wang (2020) about Kumaraswamy distribution and Weibull distribution by saying that they are the GB1 distribution at $\alpha = 1$: $GeneralizedBetaDistributionI[1, \infty, \gamma, 1] \sim WeibullDistribution[\gamma, 1] \sim KumaraswamyDistribution[\gamma, \infty]$.

Similar to the Equations (6)-(9) of the shape factor of GB1 at the boundary of the parameters space of α and γ, we can get the *S* and *K* boundary values as well from Equation (11) and (12):

$$\lim_{\alpha\to 0}\lim_{\beta\to\infty} S \sim \frac{Gamma[\alpha]^{\frac{1}{2}}\, Gamma\left[\frac{3}{\gamma}\right]}{Gamma\left[\frac{2}{\gamma}\right]^{\frac{3}{2}}}, \lim_{\alpha\to 0}\lim_{\beta\to\infty} K \sim \frac{Gamma[\alpha]\, Gamma\left[\frac{4}{\gamma}\right]}{Gamma\left[\frac{2}{\gamma}\right]^{2}} \quad ...(14)$$

$$\lim_{\gamma\to\infty}\lim_{\beta\to\infty} S = \frac{PolyGamma[2, \alpha]}{PolyGamma[1, \alpha]^{3/2}}, \lim_{\gamma\to\infty}\lim_{\beta\to\infty} K = 3 + \frac{PolyGamma[3, \alpha]}{PolyGamma[1, \alpha]^{2}} \quad ...(15)$$

$$\lim_{\gamma\to 0}\lim_{\beta\to\infty} S \sim \gamma^{\frac{\alpha}{2}-\frac{1}{4}}\left(\frac{27}{8}\right)^{\frac{1}{\gamma}}\left(\frac{9}{8}\right)^{\frac{\alpha}{2}}\left(\frac{2Gamma[\alpha]}{3}\right)^{\frac{1}{2}}\pi^{-\frac{1}{4}}, \lim_{\gamma\to 0}\lim_{\beta\to\infty} K \sim \frac{\gamma^{\alpha-\frac{1}{2}}16^{\frac{1}{\gamma}}\, Gamma[\alpha]}{\sqrt{2\pi}} \quad ...(16)$$

$$\lim_{\alpha\to\infty}\lim_{\beta\to\infty} S \sim \frac{3-\gamma}{\gamma\sqrt{\alpha}}, \lim_{\alpha\to\infty}\lim_{\beta\to\infty} K \sim 3 \quad ...(17)$$

From Equation (7) and (15), at the upper boundary $\gamma \to \infty$, when α increase from 0 to ∞, GB1 S monotonically increase from -2 to 0, *K* monotonically decrease from 9 to 3, and SF monotonically increase from 9/4 to ∞. These are exactly formulas for $ExpGammaDistribution[\alpha, \theta, \mu]$ (called Log gamma in McDonald *et al.* (2011)), i.e., $GeneralizedBetaDistributionI[\alpha, \infty, \infty, 1] \sim ExpGammaDistribution[\alpha, \theta, \mu]$.

From Equation (9) and (17), at the right boundary $\alpha \to \infty$, GB1 $S \to 0$, $K \to 3$, and $SF \to \infty$, i.e., $GeneralizedBetaDistributionI[\infty, \infty, \gamma, 1] \sim NormalDistribution[\mu, \sigma]$.

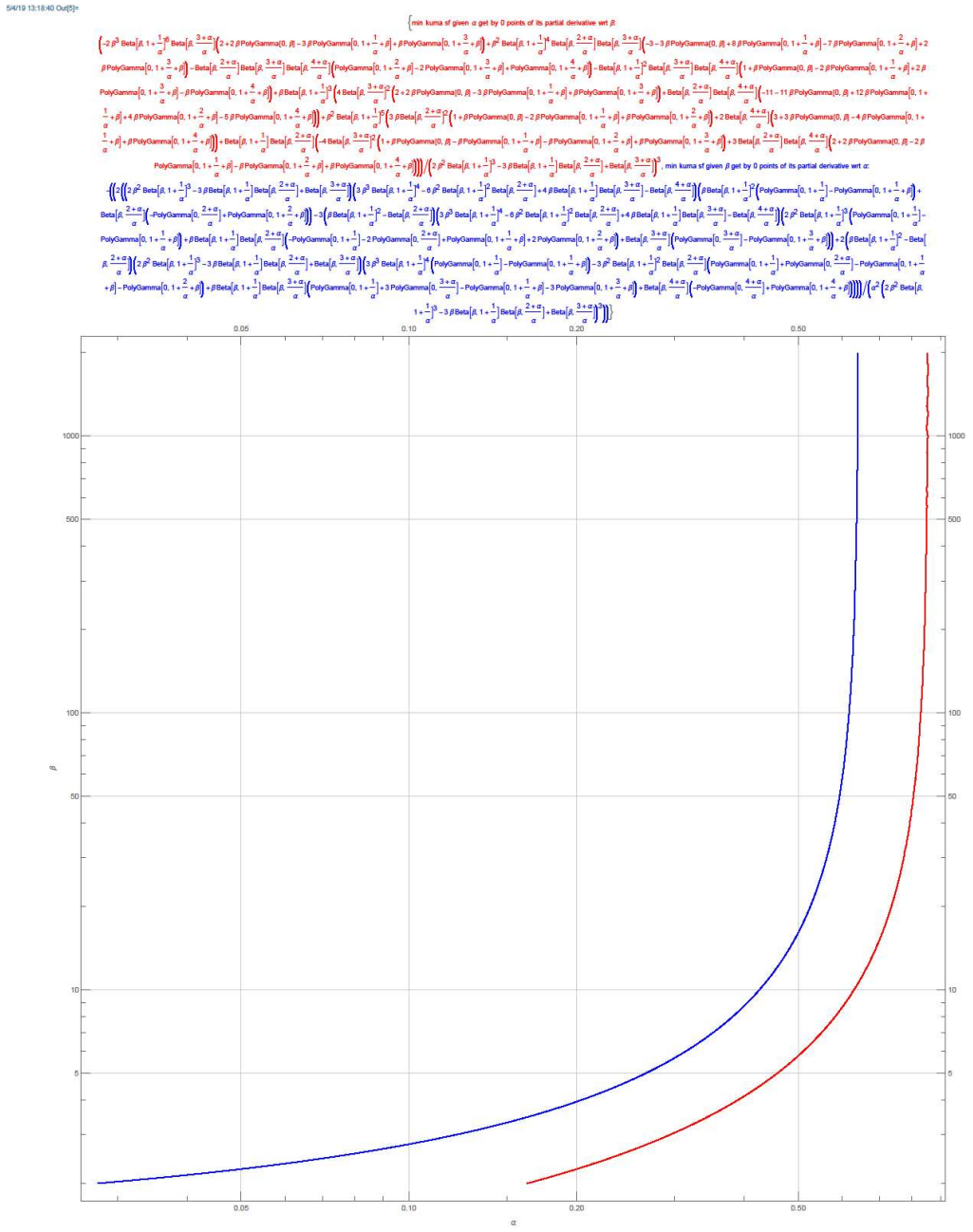
From Equation (8) and (16), at the lower boundary $\gamma \to 0$, GB1 $S \to \infty$, $K \to \infty$, and $SF \to \infty$.

From Equation (6) and (14), at the left boundary $\alpha \to 0$, GB1 $S \to \infty$, $K \to \infty$, and *SF* monotonically decrease from ∞ to 1.125 when γ increase from 0 to ∞. GB1 *SF* is discontinuous at the top left corner of the α and γ parameters space.

SF at the left boundary $\alpha \to 0$ attain minimum at $\gamma \to \infty$, and at the top boundary $\gamma \to \infty$ attain minimum at $\alpha \to 0$: the top left corner is the SF minimum point. From Wang (2020) Figure 13 and similar plots when zoom in to higher γ portion (which is also plots for GB1 since GB2 and GB1 are asymptotic equivalent), we see that the conditional minimum point for given α converge to the top left corner, in GB1. In the positive skewness region, the conditional minimum point for given γ converge to the left boundary, and then converge to the top left corner along the left boundary. In the negative skewness region, the conditional minimum point for given γ converge to the top left corner directly.

This conditional minimum point converge to boundary minimum point is also seen in distribution where the boundary shape factor has interior minimum point. For example, the Kumaraswamy distribution (a sub class of GB1) and its upper boundary asymptotic equivalent Weibull distribution have similar plot in Figure 5, where the horizontal minimum SF points converge to $\alpha = 0.6411$, the minimum *SF* point of the Weibull distribution at the upper boundary (Wang (2020) Figure 7 in section 5.3 and section 5.1 for SF values discussion). The vertical minimum *SF* points converge to $\alpha = 0.8552$

Figure 5: Kumaraswamy Distribution SF Derivatives D[SF, β] (in Red) and D[SFB, α] (in Blue) 0 Contour Plot

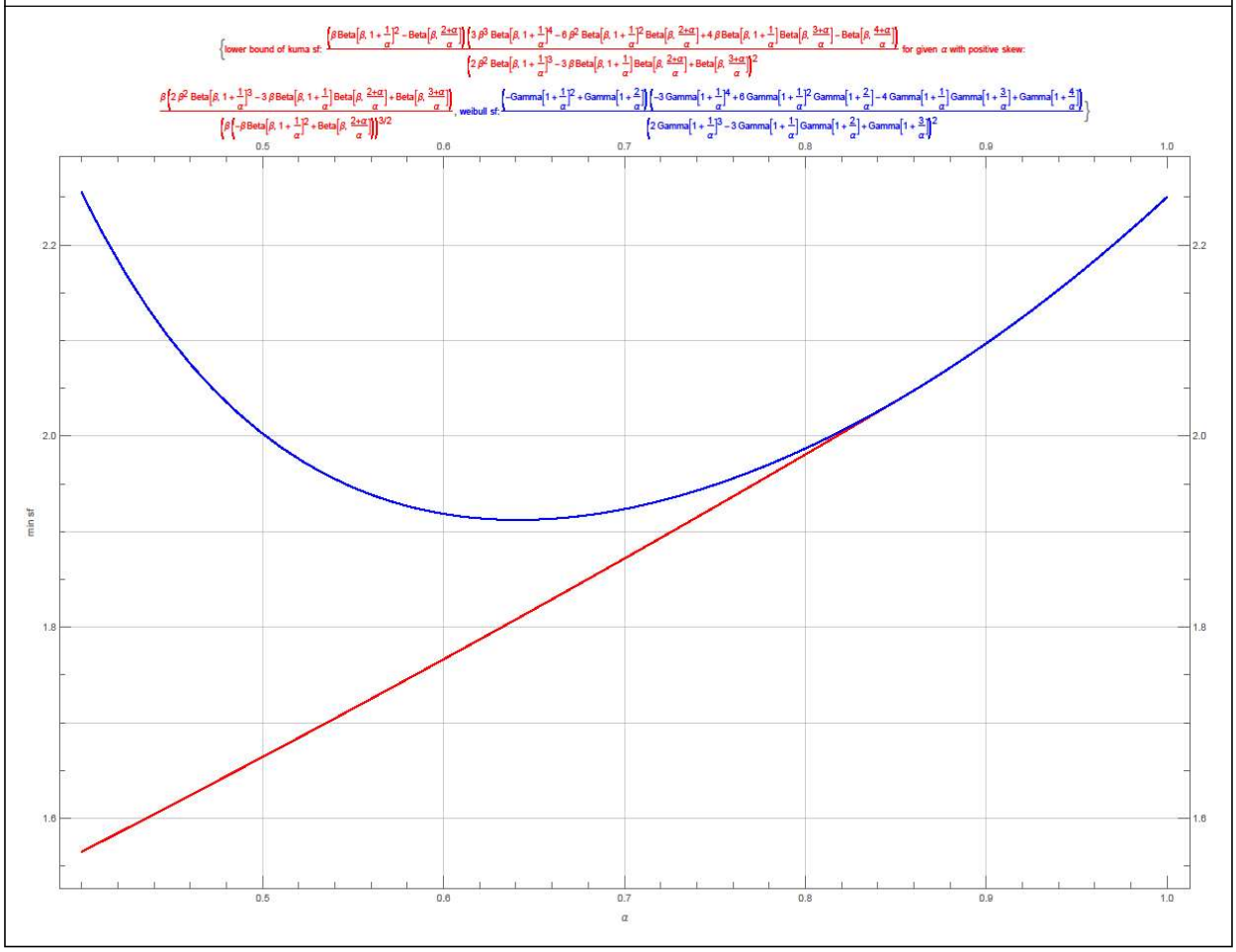


at the upper boundary. When $\alpha < 0.8552$, the vertical minimum SF attains at finite β. When $\alpha > 0.8552$, the vertical minimum *SF* attains at the upper boundary where $\beta = \infty$, which is obvious from the contour plot of the β derivative of *SF*, or from the vertical minimum *SF* (Wang (2020) Figure 8) and the Weibull distribution SF plot together, Figure 6.

GB1 *SF* is continuous at the top right and the bottom left corner of the α and γ parameters space, which always converges to ∞ no matter what route or contour we take to approach these corners. We will see later in section 2.2.1 that similar to the top left corner, the bottom right corner is also a singular point, in the sense that when we approach it from different contours we can get different (including both finite and infinite) limit values. This contour limit or contour analysis can be viewed as a generalization of the asymptotic analysis, which simply used vertical or horizontal lines as contours.

In deducting these limit formulas, we see that some obvious asymptotic order comparison of the terms in the *S*, *K*, *SF* formulas can be slightly generalized, which may be of interest by them own. When all other variables are positive and finite:

Figure 6: Kumaraswamy Distribution Vertical Minimum SF (in Red) and Weibull Distribution SF Plot (in Blue)



$$\lim_{\alpha\to\infty}\frac{Gamma[\alpha+a]^m\,Gamma[\alpha+b]^n}{Gamma\left[\alpha+\frac{ma+nb}{m+n}\right]^{m+n}}=1 \qquad ...(18)$$

$$\lim_{\gamma\to 0}\frac{Gamma\left[\alpha+\frac{a}{\gamma}\right]^m\,Gamma\left[\alpha+\frac{b}{\gamma}\right]^n}{Gamma\left[\alpha+\frac{ma+nb}{\gamma(m+n)}\right]^{m+n}}\sim\left(\frac{a^m b^n}{\left(\frac{ma+nb}{m+n}\right)^{m+n}}\right)^{\alpha-\frac{1}{2}}\left(\frac{a^{ma}b^{nb}}{\left(\frac{ma+nb}{m+n}\right)^{ma+nb}}\right)^{\frac{1}{\gamma}} \qquad ...(19)$$

$$\frac{a^m b^n}{\left(\frac{ma+nb}{m+n}\right)^{m+n}}<1 \ \textit{since log x is strictly cancave} \qquad ...(20)$$

$$\frac{a^{ma}b^{nb}}{\left(\frac{ma+nb}{m+n}\right)^{ma+nb}}>1 \ \textit{since x log x is strictly canvex} \qquad ...(21)$$

Equation (18) says that when $\alpha\to\infty$ the product factors in GB1 *SF* converge to 0 and are amenable to the derivative of log function trick (Wang (2020) Equation (12)) for calculating asymptotic power order and finally give us the asymptotic Equations (9) and (17).

Equation (19) says that when $\gamma\to 0$ the terms in the summation expression inside the product factors in GB1 *SF* have dominant terms that have the largest power order coefficients of $1/\gamma$, all other terms can be ignored. Together with the simplification by Equation (2), we then get Equations (8) and (16).

2.1.3. Minimum Shape Factor for Given Skewness

As discussed in Wang (2020) section 3.1, we can utilize the ratio of the tangent of the contour of the skewness and the tangent of the contour of the kurtosis. The shape factor for given skewness has interior minimum/maximum in the parameters space only if the ratio of the tangent equals 1 (alternatively the differences of the tangent equals 0), or geometrically the contours tangent to each other at that point. Conversely, the minimum/maximum shape factor will only be attained at the parameters space boundaries. An alternative way to check this is by examining the intersections of the skewness and kurtosis contours, interior extreme value cannot exist if we cannot find a pair of contours from the two families that have two intersections.

The ratio contour plot is in Figure 7 and the intersection plot is in Figure 8.

From Figures 7 and 8 and similar plots zooming to different regions of the parameters space that the ratio will never be 1 and we always see 1 intersection point, we can conclude that the shape factor condition on given skewness will have extreme values on the parameters boundary. The maximum will be on the top right, bottom right, or right boundary where $\alpha \to \infty$, and the minimum will be on the top left boundary where $\gamma \to \infty$ and $\alpha \to 0$. To confirm this prediction, we plot the contour of SF against S and α in Figure 9, which shows clearly the monotonically increasing of SF vs α for given S along the horizontal lines. Combining with the plots in Figures 1 and 2, we see that for given positive skewness,

Figure 7: GB1 Distribution Ratio of Tangent of Skewness Contour to Kurtosis Contour Plot

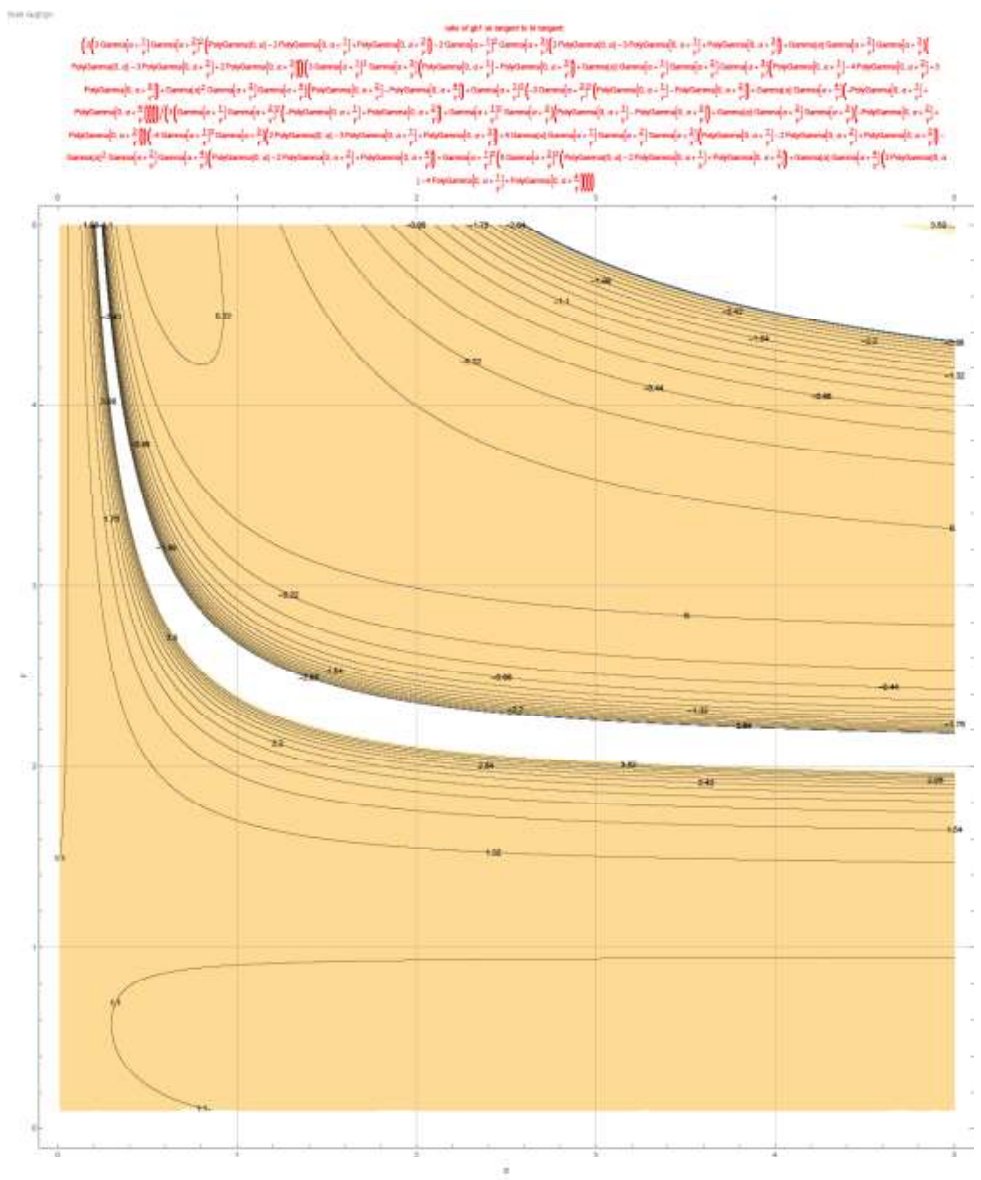
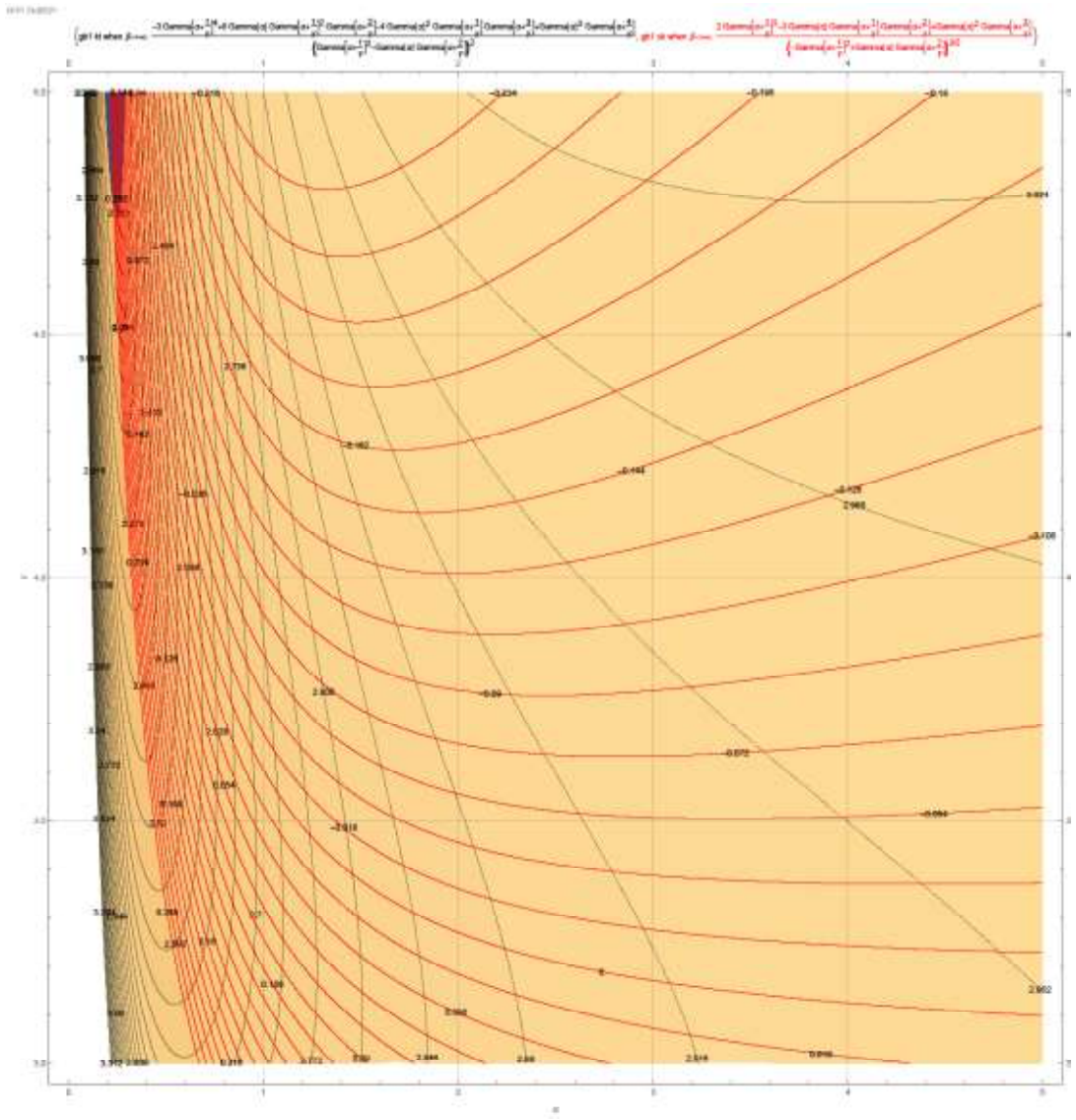


Figure 8: GB1 Distribution Skewness Contour (in Red) and Kurtosis Contour Plots



the minimum *SF* will be on the top left corner and the maximum *SF* will be on the bottom right corner. We will study these corner limits in the next section.

2.2. GB1 contour analysis part one

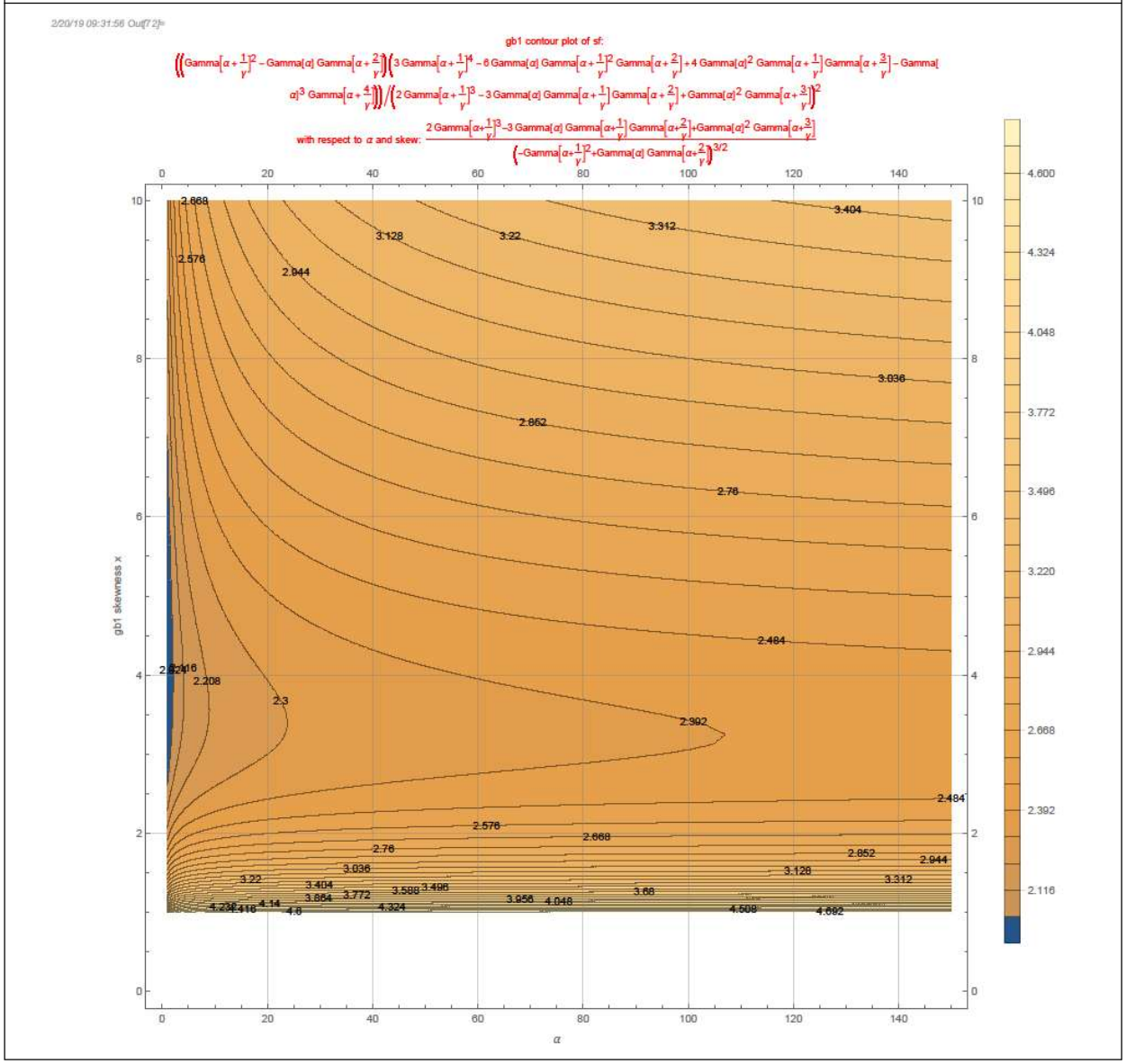
2.2.1. GB1 $\beta \to \infty$, $\alpha \to \infty$, and $\gamma \to 0$

As maintained in the last section, the bottom right corner of the α and γ parameters space of GB1 is a singular point, we need to use contour analysis (contour limit and contour minimum or maximum) to study its *S*, *K*, and *SF* properties. The simplest form of the contour is via power function: $\gamma = k/\alpha^a$. Symbolic limit using general *a* with asymptotic expansion of the Gamma function used in *S*, *K*, and *SF* at infinite do not work. This difficult can be addressed by experiments with specific values of *a*. It is found that when $a > 1/2$, the terms in the sum expression inside the product factors of *SF* have dominant term; and when $a < 1/2$, the terms are asymptotic equivalent; but when $a = 1/2$, the terms are asymptotic proportional to each other with their limit ratio not necessarily 1.

When $a = 1$, we can utilize the following heuristic trick:

$$if \ \frac{\frac{d\log\left(f\left(x\right)\right)}{dx}}{x} \to A, \ then \ possibly \ f\left(x\right) \sim x^B e^{Ax}, \ where \ x \to \infty \qquad ...(22)$$

Figure 9: GB1 Distribution SF Contour Against Skewness and α Plot



We then get the asymptotic formula for contour $\gamma = k/\alpha$:

$$\lim_{\alpha\to\infty}\lim_{\beta\to\infty} S \sim \frac{e^{\frac{\alpha\left(k\,Log[k]-3(2+k)\,Log[2+k]+2(3+k)\,Log[3+k]\right)}{2k}}(2+k)^{3/4}}{k^{1/4}\sqrt{3+k}} \quad ...(23)$$

$$\lim_{\alpha\to\infty}\lim_{\beta\to\infty} K \sim \frac{e^{\frac{\alpha\left(k\,Log[k]-2(2+k)\,Log[2+k]+(4+k)\,Log[4+k]\right)}{k}}(2+k)}{\sqrt{k(4+k)}} \quad ...(24)$$

$$\lim_{\alpha\to\infty}\lim_{\beta\to\infty} SF \sim \frac{e^{\frac{\alpha\left((2+k)\,Log[2+k]-2(3+k)\,Log[3+k]+(4+k)\,Log[4+k]\right)}{k}}(3+k)}{\sqrt{(2+k)(4+k)}} \quad ...(25)$$

The coefficients of α in Equations (23)-(25) are all positive when k is positive, therefore $S, K, SF \to \infty$. This is also true when taking symbolic limit for contours with $1/2<a<1$ or $a>1$.

At the critical value $a = 1/2$ we get the asymptotic formula for contour $\gamma = k/\alpha^{\frac{1}{2}}$:

$$\lim_{\alpha\to\infty}\lim_{\beta\to\infty}\{S,K,SF\}\sim\left\{\sqrt{-1+e^{\frac{1}{k^2}}}\left(2+e^{\frac{1}{k^2}}\right),\ -3+3e^{\frac{2}{k^2}}+2e^{\frac{3}{k^2}}+e^{\frac{4}{k^2}},\ \frac{-3+e^{\frac{2}{k^2}}\left(3+e^{\frac{1}{k^2}}\left(2+e^{\frac{1}{k^2}}\right)\right)}{\left(-1+e^{\frac{1}{k^2}}\right)\left(2+e^{\frac{1}{k^2}}\right)^2}\right\} \quad ...(26)$$

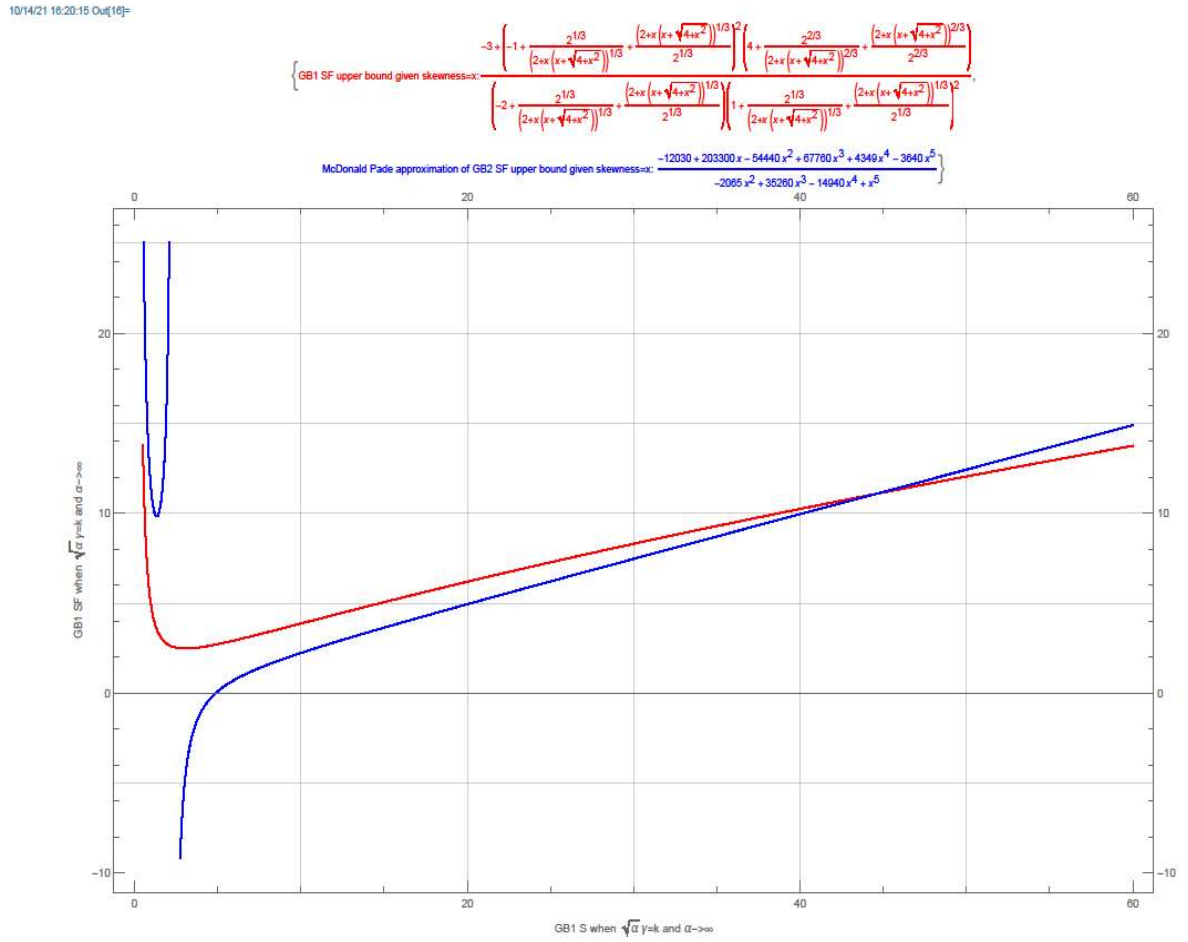
The Equation (26) is also the formula for $LogNormalDistribution[\mu,\ 1/k]$, i.e., when $\alpha\to\infty$, $GeneralizedBetaDistributionI[\alpha,\ \infty,\ k/\sqrt{\alpha},\ 1]\sim LogNormalDistribution[\mu,\ 1/k]$. This contour limit gives the upper bound of the shape factor for given positive skewness for GB1 and GG distribution that is lower than the upper bound of GB2 (McDonald *et al.* (2011) Figures 1 and 2). Solve (26) we get:

when

$$s=x>0,\ SF=\frac{-3+\left(-1+\frac{2^{1/3}}{\left(2+x\left(x+\sqrt{4+x^2}\right)\right)^{1/3}}+\frac{\left(2+x\left(x+\sqrt{4+x^2}\right)\right)^{1/3}}{2^{1/3}}\right)^2\left(4+\frac{2^{2/3}}{\left(2+x\left(x+\sqrt{4+x^2}\right)\right)^{2/3}}+\frac{\left(2+x\left(x+\sqrt{4+x^2}\right)\right)^{2/3}}{2^{2/3}}\right)}{\left(-2+\frac{2^{1/3}}{\left(2+x\left(x+\sqrt{4+x^2}\right)\right)^{1/3}}+\frac{\left(2+x\left(x+\sqrt{4+x^2}\right)\right)^{1/3}}{2^{1/3}}\right)\left(1+\frac{2^{1/3}}{\left(2+x\left(x+\sqrt{4+x^2}\right)\right)^{1/3}}+\frac{\left(2+x\left(x+\sqrt{4+x^2}\right)\right)^{1/3}}{2^{1/3}}\right)^2} \quad ...(27)$$

The plot of (27) together with the Pade-approximation of the GB2 upper bound (McDonald *et al.* (2011) appendix) are in Figure 10.

Figure 10: GB1 (in Red) and GB2 (in Blue) Distribution Maximum Shape Factor with Respect to Skewness Plot



Along the contour $\gamma = k/\alpha^a$ when a<1/2, we cannot obtain general symbolic limit results. But for a = 1/4 or 1/3, we will get $S \to -\infty$, $K \to \infty$, and $SF \to 4/3$, by using the second order asymptotic expansion of the product Gamma function $\frac{e^z Gamma[z]}{\sqrt{2\pi} z^{-\frac{1}{2}+z}}$ at infinite, and using the following simplification (the condition in it is a clue for why ½ is the critical value):

$$(s+l)^{-o+n(s+l)} \sim e^{sn} l^{-o+n(s+l)} \text{ when } \frac{s}{l} \to 0 \text{ and } \frac{s^2}{l} \to 0 \quad ...(28)$$

The Equation (28) is an example where simplification can bring about as many problems as producing conclusions, if not more. Without using (28), for a = 1/4 or 1/3, and using the first, second, third, or even fourth order expansion of $\frac{e^z Gamma[z]}{\sqrt{2\pi} z^{-\frac{1}{2}+z}}$, we get $SF \to \infty$ if we take the symbolic limit directly.

In a different manipulation without using (28), but using second order expansion of $\frac{Gamma[w]^h Gamma[x]}{Gamma[z]^k}$ so that the exponential and higher order factors will cancel out and the remaining part are only power functions, we get $SF \to \infty$ when a<1/4, $SF \to$ 4/3+k^4/4608 when a = 1/4, $SF \to$ 4/3 when 1/4<a<1/2. But if we use third order expansion of $\frac{Gamma[w]^h Gamma[x]}{Gamma[z]^k}$, we get $SF \to \infty$ when a<1/4, $SF \to$ 4/3+59 k^4/69120 when a = 1/4, $SF \to$ 4/3 when 1/4<a<1/2.

Whether it is 4/3 or it is ∞, it is the problem.

For more exploration, using first order expansion of $\frac{e^z Gamma[z]}{\sqrt{2\pi} z^{-\frac{1}{2}+z}}$ and (28), and using Wang (2020) Equation (12)'s calculation of limit of x*dlogf(x)/dx to estimate the power function power order, and followed by symbolic quotient calculation for estimating the coefficients, we get $SF \sim$ 5 k^3 α^(2-3a) when a<1/4. But if we use the second order expansion of $\frac{e^z Gamma[z]}{\sqrt{2\pi} z^{-\frac{1}{2}+z}}$ instead, we get SF ~ 16 α^(–2+a)/(243k), which unlike the first order expansion conclusion, is obviously wrong since it converges to 0 when a<2.

The final resolution of all these confusions is through induction with sample values of a and k, and through the following observation: when $x \to \infty$ and f(x) $\to \infty$, the usual infiniteness magnitude or form of f(x) is either logarithmic, power, exponential, or exponential of power. We can then write f(x) ~ log(x) ^ G * x ^ B * ℮ ^ (A * x ^ K). And we can use the limit of log(log(f(x)))/log(x) to estimate K, the limit of log(f(x)) /x^K to estimate A, the limit of log(f(x))/log(x) to estimate B when K<=1, lastly the limit of log(f(x))/log(log(x)) to estimate G when K<=0 and B = 0. Using derivatives such as x*d(log(f(x)))/dx may also be able to estimate B, and using d(log(f(x)))/dx may be able to estimate A when K = 1, but they do not work in our case. And in cases both methods work, they are slower than the log(f(x))/log(x) or log(f(x))/x limits method.

We first fix a = 1/4, and for various values of k, take the symbolic limit of

$$Log\left[\left(Gamma\left[\alpha+\frac{\alpha^a}{k}\right]^2 - Gamma[\alpha]Gamma\left[\alpha+\frac{2\alpha^a}{k}\right]\right)\right.$$

$$\left(3Gamma\left[\alpha+\frac{\alpha^a}{k}\right]^4 - 6Gamma[\alpha]Gamma\left[\alpha+\frac{\alpha^a}{k}\right]^2 Gamma\left[\alpha+\frac{2\alpha^a}{k}\right]\right.$$

$$\left.+4Gamma[\alpha]^2 Gamma\left[\alpha+\frac{\alpha^a}{k}\right]Gamma\left[\alpha+\frac{3\alpha^a}{k}\right] - Gamma[\alpha]^3 Gamma\left[\alpha+\frac{4\alpha^a}{k}\right]\right)$$

$$/\left(2Gamma\left[\alpha+\frac{\alpha^a}{k}\right]^3 - 3Gamma[\alpha]Gamma\left[\alpha+\frac{\alpha^a}{k}\right]Gamma\left[\alpha+\frac{2\alpha^a}{k}\right]\right.$$

$$\left.\left.+Gamma[\alpha]^2 Gamma\left[\alpha+\frac{3\alpha^a}{k}\right]\right)^2 //.\{a \to 1/4, k \to 2\} // Evaluate\right] / Log[\alpha]$$

i.e., using the SF formula without any series expansion. We find that the limit is always 1/2 no matter what values of the k we take. Henceforth we fix $k = 1$ in the explorations for other values of a. For various values of a in the interval [0, 1/2], we calculate all their power orders, and find that it is of the form $1–2a$. Similar experiments are done for other intervals. The results are summarized in Equation (29).

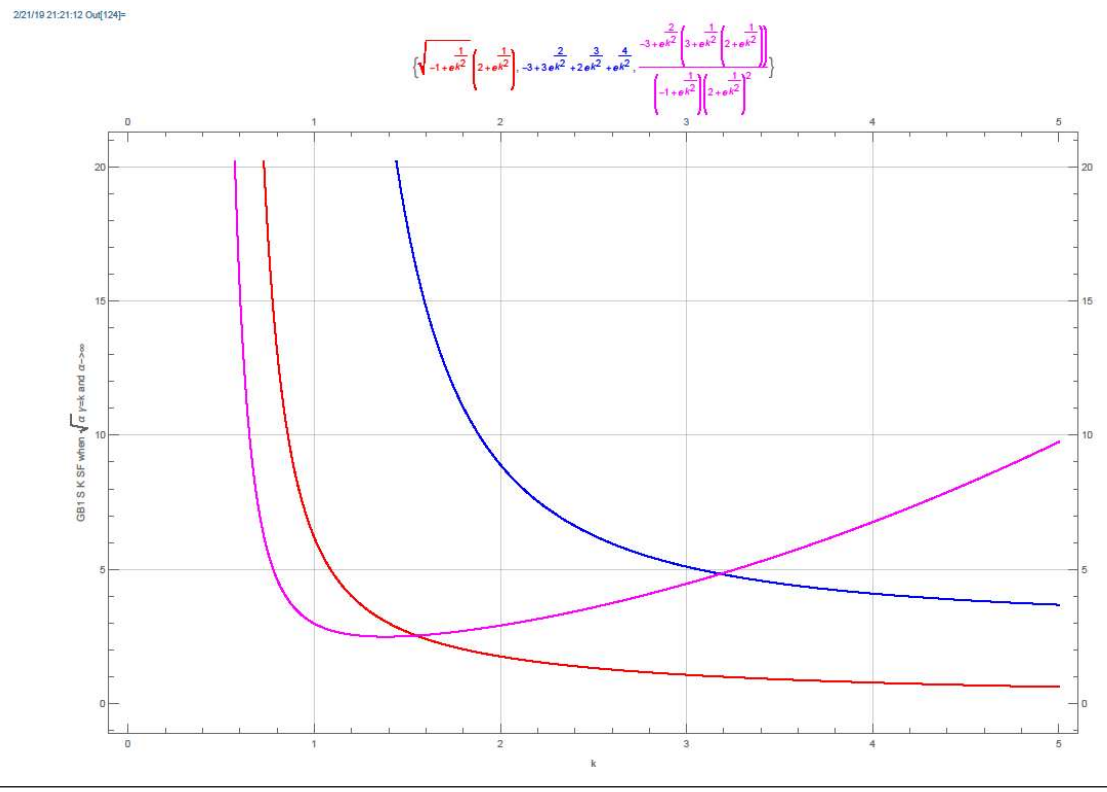
$$SF \to \begin{cases} \alpha & -1 \le a \le 0 \\ \alpha^{1-2a} & 0 < a < \frac{1}{2} \\ \frac{-3+3e^2+2e^3+e^4}{(-1+e)(2+e)^2} & a == \frac{1}{2} \\ e^{\alpha^{-1+2a}} & \frac{1}{2} < a < 1 \\ \left(\frac{84375}{65536}\right)^{\alpha} & a == 1 \\ \left(\frac{1024}{729}\right)^{\alpha^a} & 1 < a \end{cases} \qquad ...(29)$$

When a<-1, such as when a = -2, we are not able to get the symbolic limit. But for a in [–1, 1/2), from Equation (29) we know $SF \to \infty$ not the 4/3, resolved our quandary. Equation (29) become (9) when $a = 0$, (25) when $a = 1$, and (26) when $a = 1/2$.

Our solution methodology is more of mixed nature from induction and deduction, not purely induction. For a purely induction way, using a sequence of selected and fixed a, we can calculate the numerical SF value for α up to 10.^5 using machine accuracy numbers, and for α up to 10^7 using arbitrary accuracy numbers, then we may want to find the tendency. But the subsequent power function fitting by NonlinearModelFit cannot find any decisive results, perhaps due to not higher enough α can be calculated. Consequently, the symbolic limit with $\alpha \to \infty$ is employed.

In summary, at the bottom right corner of the α and γ parameters space of GB1, when we approach it along different contours of $\gamma = k / \alpha^a$ for different a, the S, K, and SF behave differently. Only at a = 1/2, the asymptotically *LogNormalDistribution*[μ, 1/k] case, do we have all finite S, K, and SF. In such a case, S and K monotonically decrease

Figure 11: GB1 S (in Red) K (in Blue) SF (in Magenta) When $\gamma = k/\sqrt{\alpha}$ and $\alpha \to \infty$ Plot



with respect to k, but SF has an interior minimum, Figure 11. The contour analysis with our induction and deduction tricks is a viable tool for singularity analysis approaching either the parameters space boundaries or corners: it can help us find formulas or patterns out of irregularities and identify various critical points.

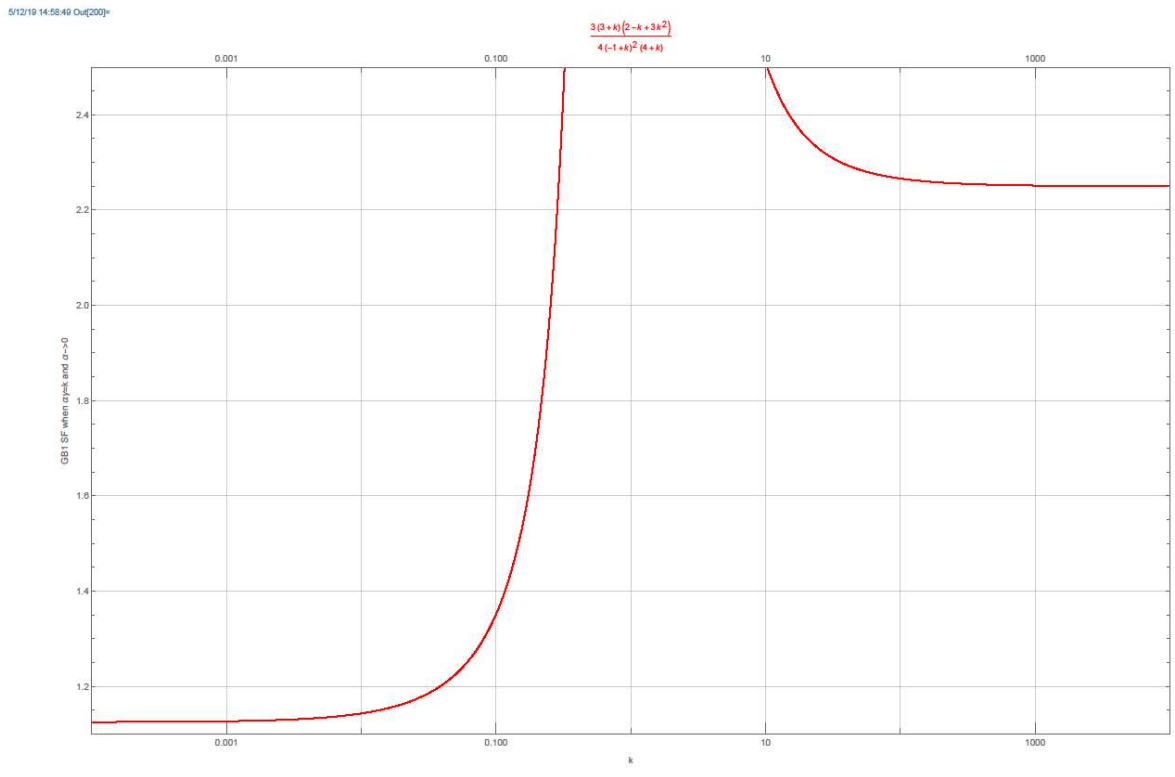
We guess why the power function form of contours have non-trivial finite limit of S, K, and SF has some relationship with the S zero contour shape. Through two points fitting we know the GB1 $S=0$ contour almost seamlessly match with the curve $\gamma = 3 + \frac{4}{9\alpha}$ when $\alpha > 100$, indicating that the S zero contour is approximately of power function form. Other distribution contour analysis can likewise get inspiration from their skewness zero-contour function form.

2.2.2. GB1 $\beta \to \infty$, $\alpha \to 0$ and $\gamma \to \infty$

The top left corner is similar to the bottom right corner, taking symbolic limit for arbitrary a would not work, but for specific values or small intervals of a, the symbolic limit will do the trick. Using the same mixed induction method, along the contour $\gamma = k / \alpha^a$, when $\alpha \to 0$ we get the following contour limit for the whole range of real a leading to left boundary or corners:

$$\{S, K, SF\} \sim \begin{cases} \{\infty, \infty, \infty\} & a < 0 \\ \left\{\infty, \infty, \dfrac{Gamma\left[\frac{2}{k}\right] Gamma\left[\frac{4}{k}\right]}{Gamma\left[\frac{3}{k}\right]^2}\right\} & a == 0 \\ \left\{\infty, \infty, \dfrac{9}{8}\right\} & 0 < a < 1 \\ \left\{-\dfrac{2(-1+k)\sqrt{\frac{2+k}{k}}}{3+k}, \dfrac{3\left(4+5k^2+3k^3\right)}{k(3+k)(4+k)}, \dfrac{3(3+k)\left(2-k+3k^2\right)}{4(-1+k)^2(4+k)}\right\} & a == 1 \\ \left\{-2, 9, \dfrac{9}{4}\right\} & 1 < a \end{cases} \quad ...(30)$$

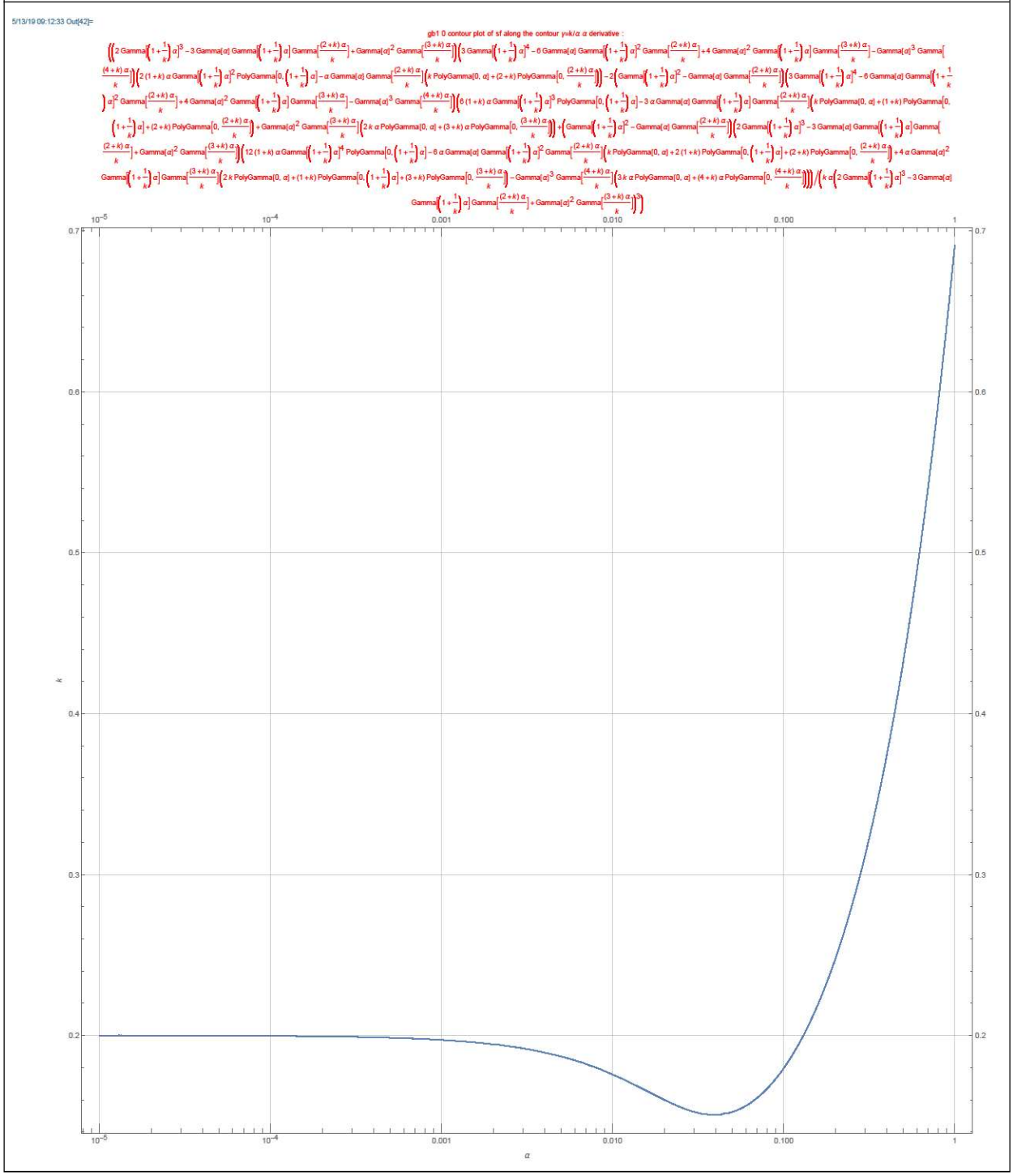
Figure 12: GB1 Shape Factor Value When $\alpha \to 0$ Along the Contour $\gamma = k/\alpha$



The piecewise function Equation (30) is bridged at the ends of the different a intervals, extrapolated with respect to the variable k whose two end points 0 or ∞ limits converge to the above or below adjacent formulas in this list. When a = 1, the limit is identical to S, K, and SF of *BetaDistribution*[k, 1], i.e., *GeneralizedBetaDistributionI*[α, ∞, k/α, 1]~*BetaDistribution*[k, 1]. This distribution is called power function or PF distribution in McDonald *et al.* (2011), where it is shown to be the lower bound of GB2 and GG. Its *SF* plot is in Figure 12.

The curve in Figure 12 is slightly higher than Wang (2018b) Figure 27; especially when k is positioned at the middle range around 1 to 10. The reason for this differences is that the contour limit is not the contour minimum, which take place in interior α values, as indicated by the zero contour plot of the α derivative of the *SF* along the contour $\gamma = k/\alpha$ in Figure 13.

Figure 13: GB1 Along $\gamma = k/\alpha$ Shape Factor α Derivative Zero Contour Plot



The *SF* at these interior minimum is in Figure 14, which is more close to Wang (2018b) Figure 27, ignoring their numerical noises produced from Mathematica FindRoot function.

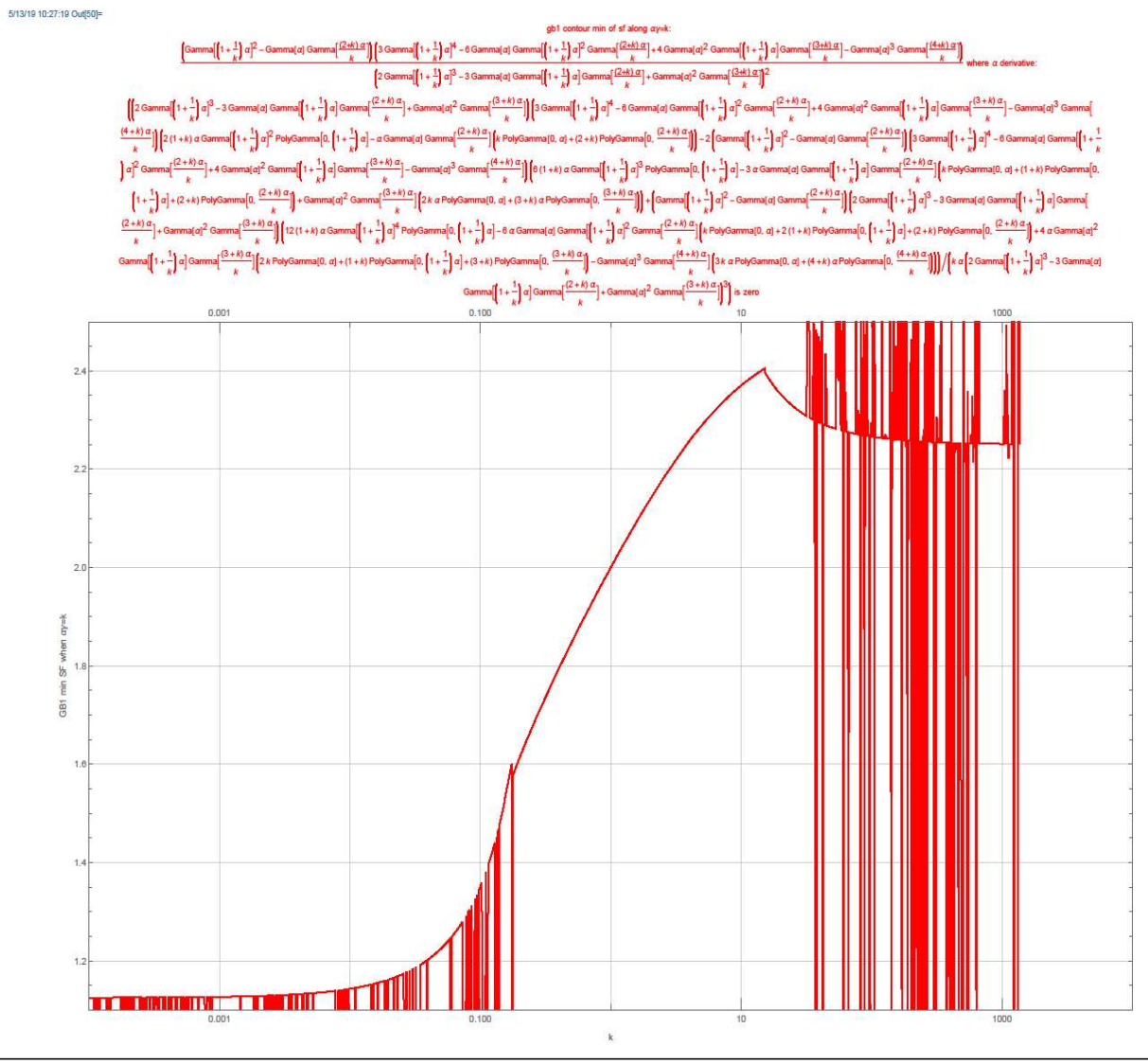
In summary, the contour analysis through contour family $\gamma = k / \alpha^a$ reveals abundant information of GB1 *SF* at the top left and the bottom right corner in the parameter space of α and γ. Some values of a, specifically 1 and 1/2 respectively, give finite *S*, *K*, and *SF*, or yield distribution asymptotic equivalent to PF or LogNormal distribution. This knowledge are of practical values for parameter range estimation or validation in distribution fitting.

2.3. GB1 Shape Factor Global Minimum

2.3.1. When $\beta \to 0$

From Figures 2-4 and Wang (2020) section 6 we know GB1, as well as GB2 and GG, have minimum SF 1.125 when $\beta \to \infty$, arrived at the top left corner. However, this is not the global minimum *SF* of GB1. If we turn to the opposite direction of $\beta \to \infty$, let $\beta \to 0$, using power series expansion of the expression $\left(\frac{Gamma[w+\beta]}{Gamma[w]}\right)^h \frac{Gamma[x]}{Gamma[x+\beta]}\left(\frac{Gamma[z]}{Gamma[z+\beta]}\right)^k$ and $\left(\frac{Gamma[w+\beta]}{Gamma[w]}\right)^h \left(\frac{Gamma[z]}{Gamma[z+\beta]}\right)^k$ at $\beta = 0$ for various combinations of w, x, z, h, and k that represent summation terms in GB1 *S*, *K*, and *SF*, we get:

Figure 14: GB1 Along $\gamma = k/\alpha$ Minimum Shape Factor Value for Given k



$$\lim_{\beta\to 0}\{S, K, SF\} \sim \left\{\infty, \infty\left(PolyGamma\left[0,\alpha\right]-2PolyGamma\left[0,\alpha+\frac{1}{\gamma}\right]+PolyGamma\left[0,\alpha+\frac{2}{\gamma}\right]\right)*\right.$$

$$\left(PolyGamma\left[0,\alpha\right]-4PolyGamma\left[0,\alpha+\frac{1}{\gamma}\right]+6PolyGamma\left[0,\alpha+\frac{2}{\gamma}\right]-4PolyGamma\left[0,\alpha+\frac{3}{\gamma}\right]+PolyGamma\left[0,\alpha+\frac{4}{\gamma}\right]\right)/$$

$$\left.\left(PolyGamma\left[0,\alpha\right]-3PolyGamma\left[0,\alpha+\frac{1}{\gamma}\right]+3PolyGamma\left[0,\alpha+\frac{2}{\gamma}\right]-PolyGamma\left[0,\alpha+\frac{3}{\gamma}\right]\right)^2\right\} \quad ...(31)$$

The Equation (31) is verified by both the first and the fourth order power expansion. The *SF* formula in Equation (31) converges to 1, the greatest lower bound of *SF*, at the bottom left corner of the parameter space of α and γ.

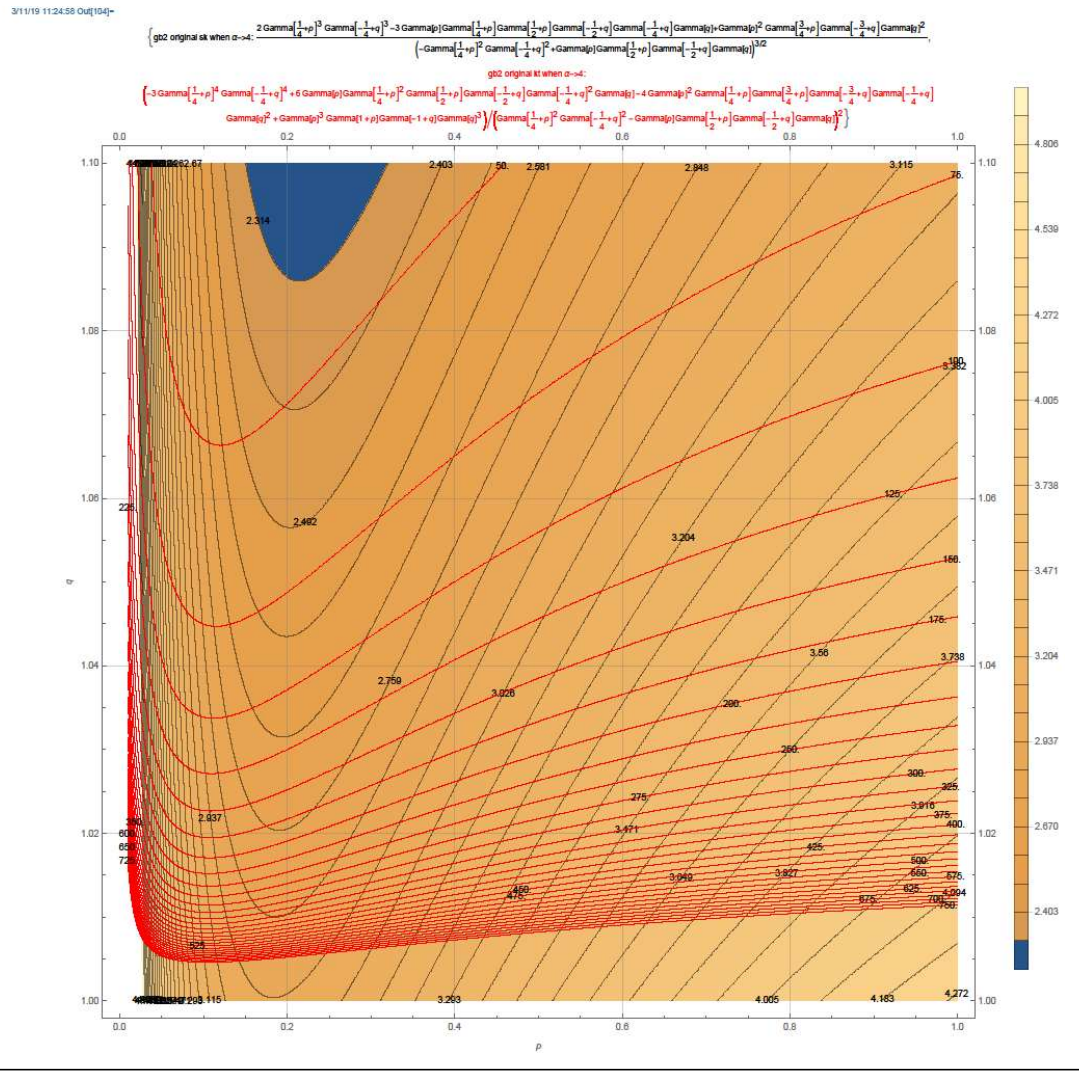
2.4. GB2 Contour Analysis

2.4.1. When $q \to 0$?

Similar to the study of GB1 for $\beta \to 0$, we want to know GB2 *SF* when $q \to 0$. Unlike GB1 whose *SF* always exist, the GB2 must have $q\alpha$>4 to guarantee SF existing. Therefore, when $q \to 0$, we need to request $\alpha \to \infty$. In this case, the asymptotic limit of $q \to 0$ while fixing α will give wrong results: the contour limit must be engaged.

To find where GB2 attain *SF* maximum, we select example value of $\alpha = 4$, explore the *S* and *SF* contour family, find that given *S*, the maximum of *SF* is at interior point of the parameter space, Figure 15.

Figure 15: GB2 S (in Black) Contours and SF (in Red) Contours Plot When $\alpha = 4$



This is confirmed by the ratio of the *S* contour tangent to the *K* contour tangent exploration. The curve where this tangent ratio equals one is interposed on the *SF* contour plot in Figure 16.

From Figure 16 we see that the maximum *SF* location of GB2 given *S* is almost a line in the *p* and *q* parameter space. Map the parameter *p* and *q* to *S* and *K* we get Figure 17.

In Figure 17, we juxtaposed our $\alpha = 4$ max S and K plot with the McDonald Pade-approximation of the GB2 empirical upper bound (McDonald *et al.* (2011) appendix), which showed curve shape similarity.

Experiment with various α: 2, 8, 16, and 32, we see that when $\alpha \to \infty$, the max *S* and *K* plot converges toward McDonald's empirical formula plot.

Using induction method on α, and guided by the numerical results of McDonald *et al.* (2011), we are lead to study on limit toward a different direction than *q*, the α direction.

Taking second order power series expansion of the expression $\left(\frac{Gamma[z+x]}{Gamma[z]}\right)^h \frac{Gamma[z+y]}{Gamma[z]}\left(\frac{Gamma[u+v]}{Gamma[u]}\right)^k \frac{Gamma[u+w]}{Gamma[u]}$ and $\left(\frac{Gamma[z+x]}{Gamma[z]}\right)^h \left(\frac{Gamma[u+v]}{Gamma[u]}\right)^k$ at x, y, v, and $w = 0$ for various combinations of z, u, x, h, y, v, k, and w that represent

Figure 16: GB2 SF Contours and the Ratio of S Tangent to K Tangent Equals One Curve When $\alpha = 4$

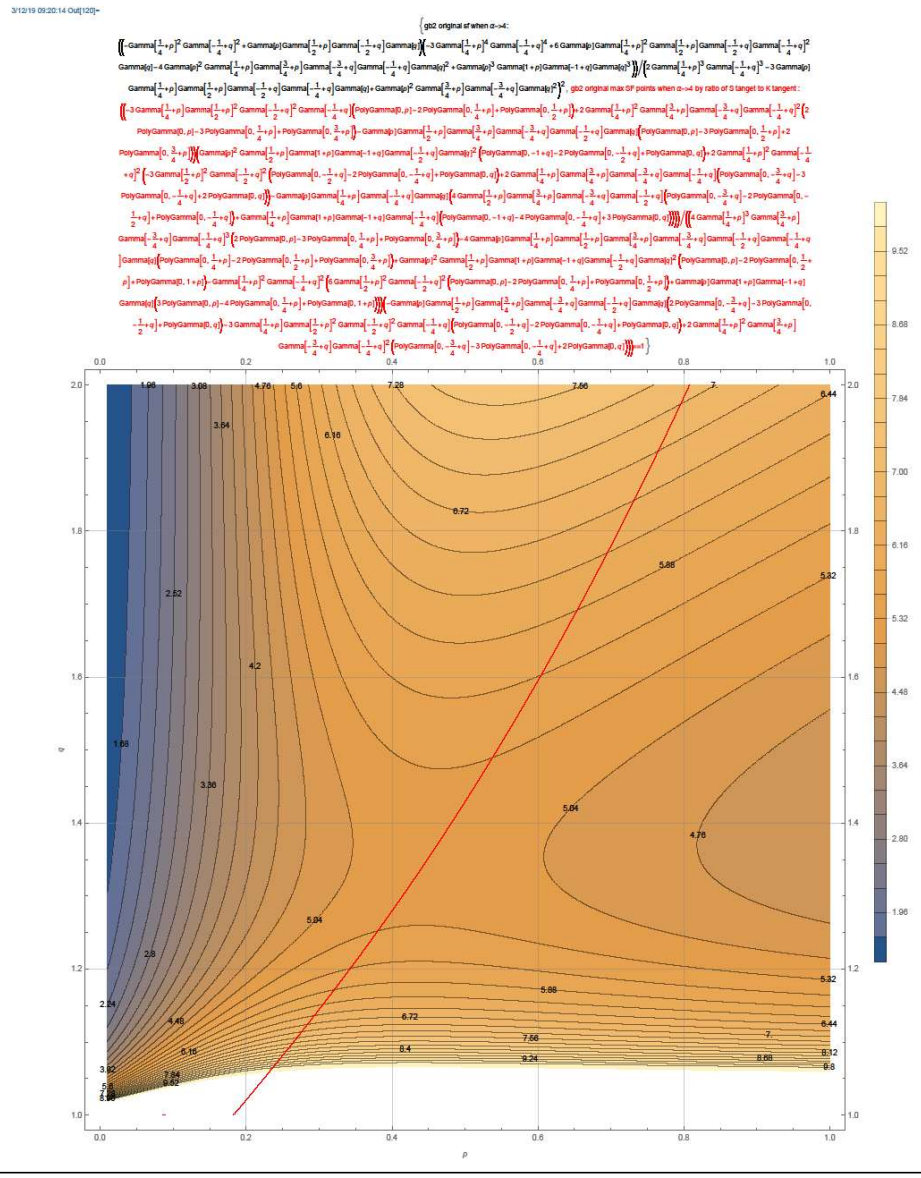
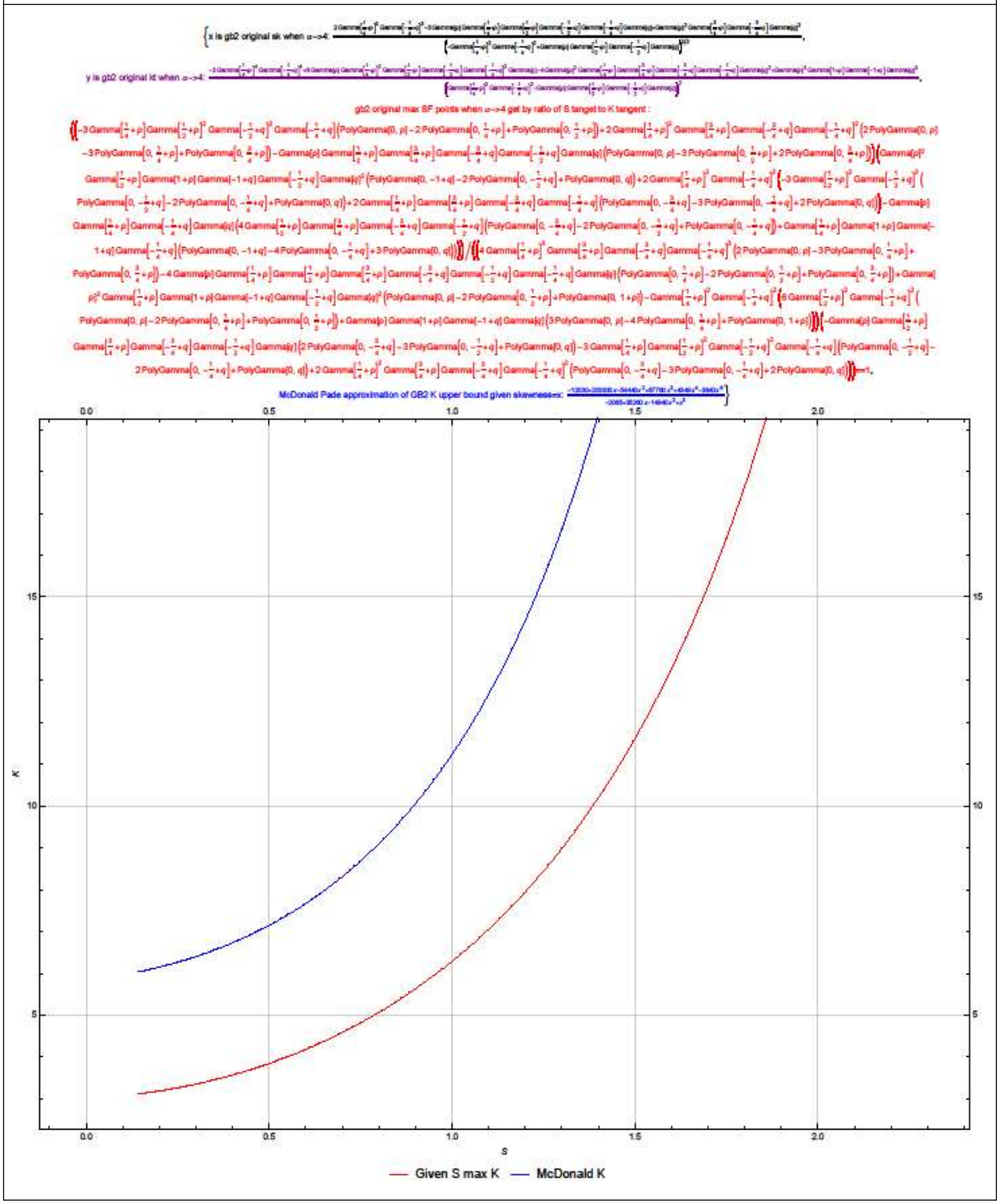


Figure 17: GB2 S and K Where the Ratio of S Tangent to K Tangent Equals One (in Red) When $\alpha = 4$

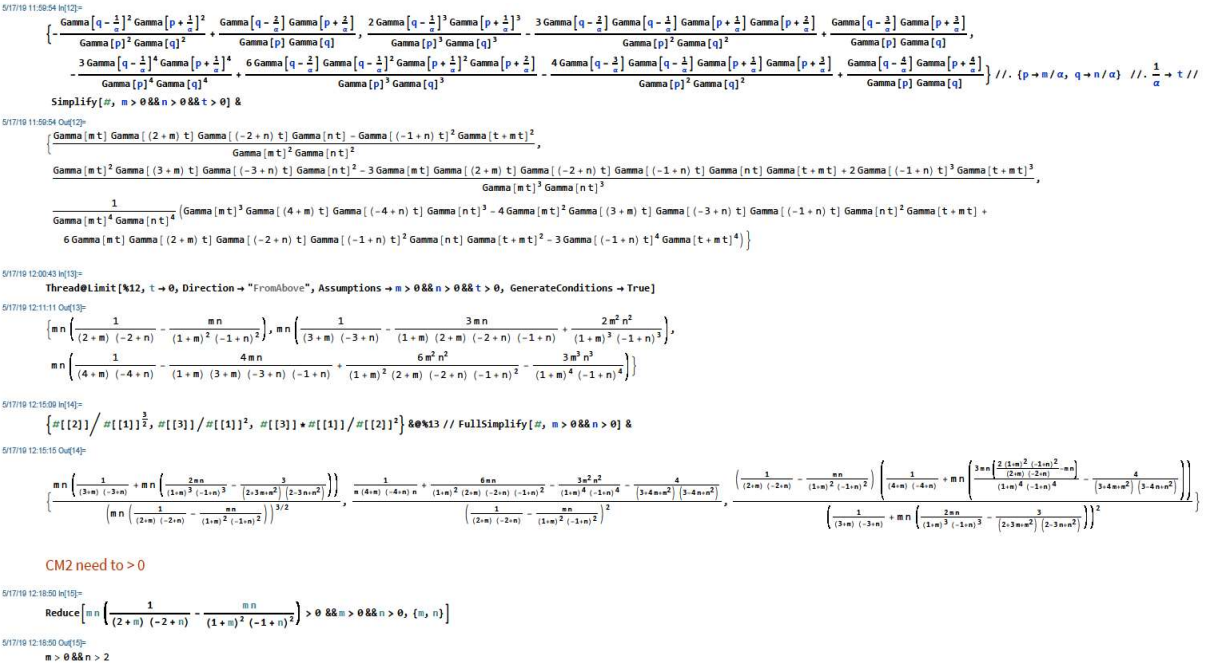


summation terms in GB2 *S*, *K*, and *SF*, we can obtain GB2 ratio of *S* tangent to *K* tangent formula and contour plot. The ratio equals one plot is just $p = q$, which hint at that the maximum *SF* given *S* is at the *p*–*q* space bottom left corner with p proportional to *q*. Therefore, we will try the simplest such contour in the next subsection.

2.4.2. When $\alpha \to \infty$ and $p = m/\alpha$, $q = n/\alpha$

Using the same second order power series expansion, but now remove the incorrect assumption that *p* and *q* is fixed, let $p = m/\alpha$, $q = n/\alpha$, and $\alpha \to \infty$, we get the contour limit formula for GB2 *S*, *K*, and *SF*, Equations (32)-(34). Later we deduct the same formula without using the series expansion, in Figure 18.

Figure 18: GB2 S, K, SF Limit Formula Deduction When $\alpha \to \infty$ and p = m/α, q = n/α



$$\lim_{\alpha\to\infty,\ p=\frac{m}{\alpha},\ q=\frac{n}{\alpha}} S \sim \frac{mn\left(\frac{1}{(3+m)(-3+n)}+mn\left(\frac{2mn}{(1+m)^3(-1+n)^3}-\frac{3}{(2+3m+m^2)(2-3n+n^2)}\right)\right)}{\left(mn\left(\frac{1}{(2+m)(-2+n)}-\frac{mn}{(1+m)^2(-1+n)^2}\right)\right)^{3/2}} \quad ...(32)$$

$$\lim_{\alpha\to\infty,\ p=\frac{m}{\alpha},\ q=\frac{n}{\alpha}} K \sim \frac{\frac{1}{m(4+m)(-4+n)n}+\frac{6mn}{(1+m)^2(2+m)(-2+n)(-1+n)^2}-\frac{3m^2n^2}{(1+m)^4(-1+n)^4}-\frac{4}{(3+4m+m^2)(3-4n+n^2)}}{\left(\frac{1}{(2+m)(-2+n)}-\frac{mn}{(1+m)^2(-1+n)^2}\right)^2} \quad ...(33)$$

$$\lim_{\alpha\to\infty,\ p=\frac{m}{\alpha},\ q=\frac{n}{\alpha}} SF \sim \frac{\left(\frac{1}{(2+m)(-2+n)}-\frac{mn}{(1+m)^2(-1+n)^2}\right)\left(\frac{1}{(4+m)(-4+n)}+mn\left(\frac{3mn\left(\frac{2(1+m)^2(-1+n)^2}{(2+m)(-2+n)}-mn\right)}{(1+m)^4(-1+n)^4}-\frac{4}{(3+4m+m^2)(3-4n+n^2)}\right)\right)}{\left(\frac{1}{(3+m)(-3+n)}+mn\left(\frac{2mn}{(1+m)^3(-1+n)^3}-\frac{3}{(2+3m+m^2)(2-3n+n^2)}\right)\right)^2}$$

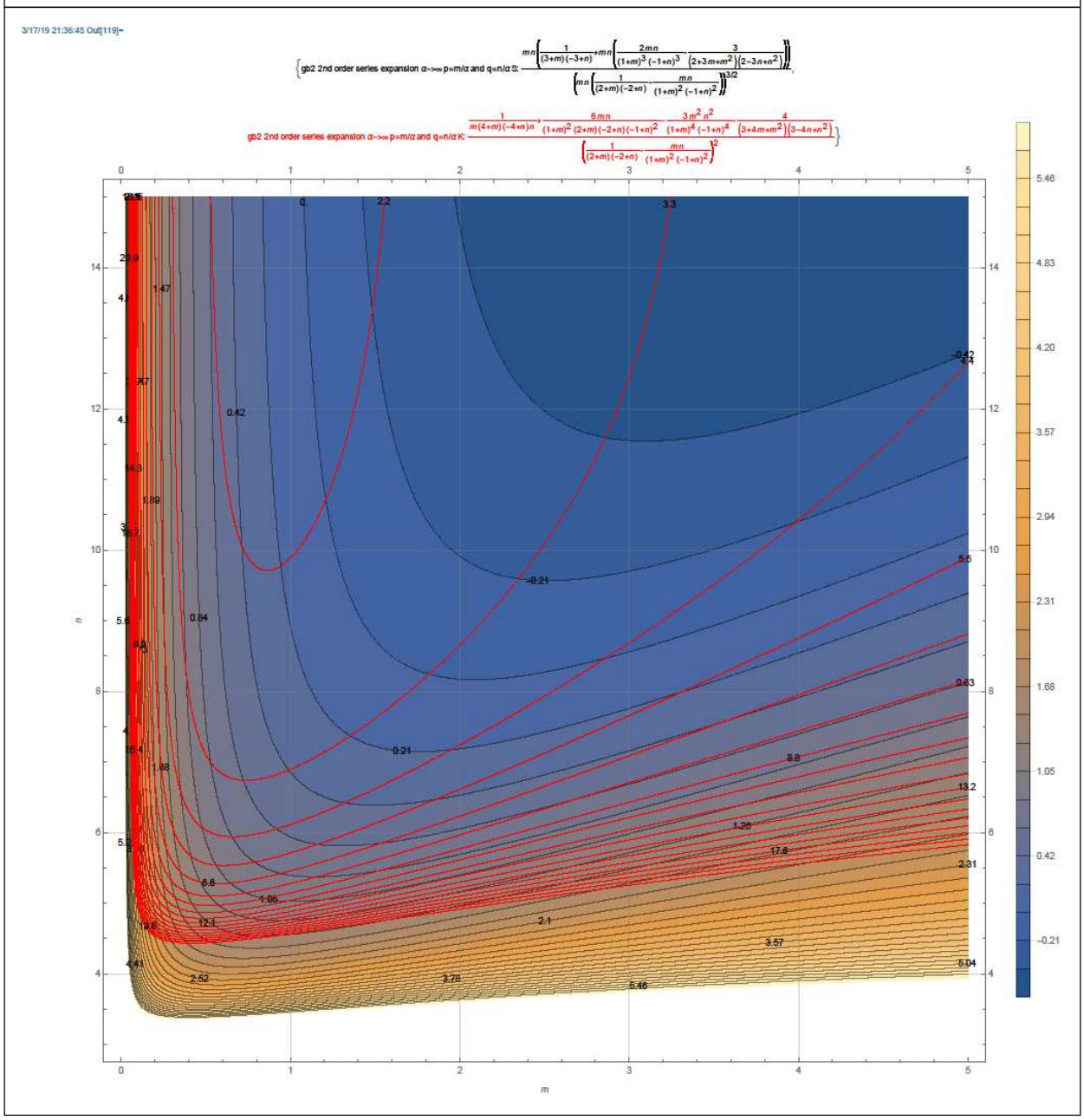
...(34)

This limit of GB2 realize maximum *SF* for given *S* at interior point of the *m*–*n* space, as seen from the contours plot of *S* and *K* that have two pairwise intersections in the bottom part in Figure 19.

We then use the ratio of the *S* contour tangent to the *K* contour tangent to locate the interior maximum position of GB2 *SF* given *S*. The ratio equals one is equivalent to the ratio expression numerator equals denominator, which in turn equivalent to the difference of the numerator with denominator equals zero. Factor out other positive parts we get the location of the *SF* maximum by Equation (35). Together with the constraint of CM[2]>0, CM[4]>0, *K*>1, and *SF*>1, we get the explicit solution of Equation (35) by Equation (36) and Equation (37). The deduction can also be done via the difference of S contour tangent to the K contour tangent equals zero route.

$$\begin{aligned} &6+23m+34m^2+24m^3+8m^4+m^5-23n+6mn+10m^2n+6m^3n+m^4n \\ &\quad +34n^2-10mn^2-24n^3+6mn^3+8n^4-mn^4-n^5=0 \end{aligned} \quad ...(35)$$

Figure 19: GB2 S Contours (in Black) and K Contours (in Red) Plot When α → ∞ and p = m/α, q = n/α



$$m > 0.6366164911641596 \;\&\&\; n == Root[-6 - 23m - 34m^2 - 24m^3 - 8m^4 - m^5 + (23 - 6m - 10m^2 - 6m^3 - m^4)\#1 + (-34 + 10m)\#1^2 + (24 - 6m)\#1^3 + (-8 + m)\#1^4 + \#1^5 \;\&, 1] \quad ...(36)$$

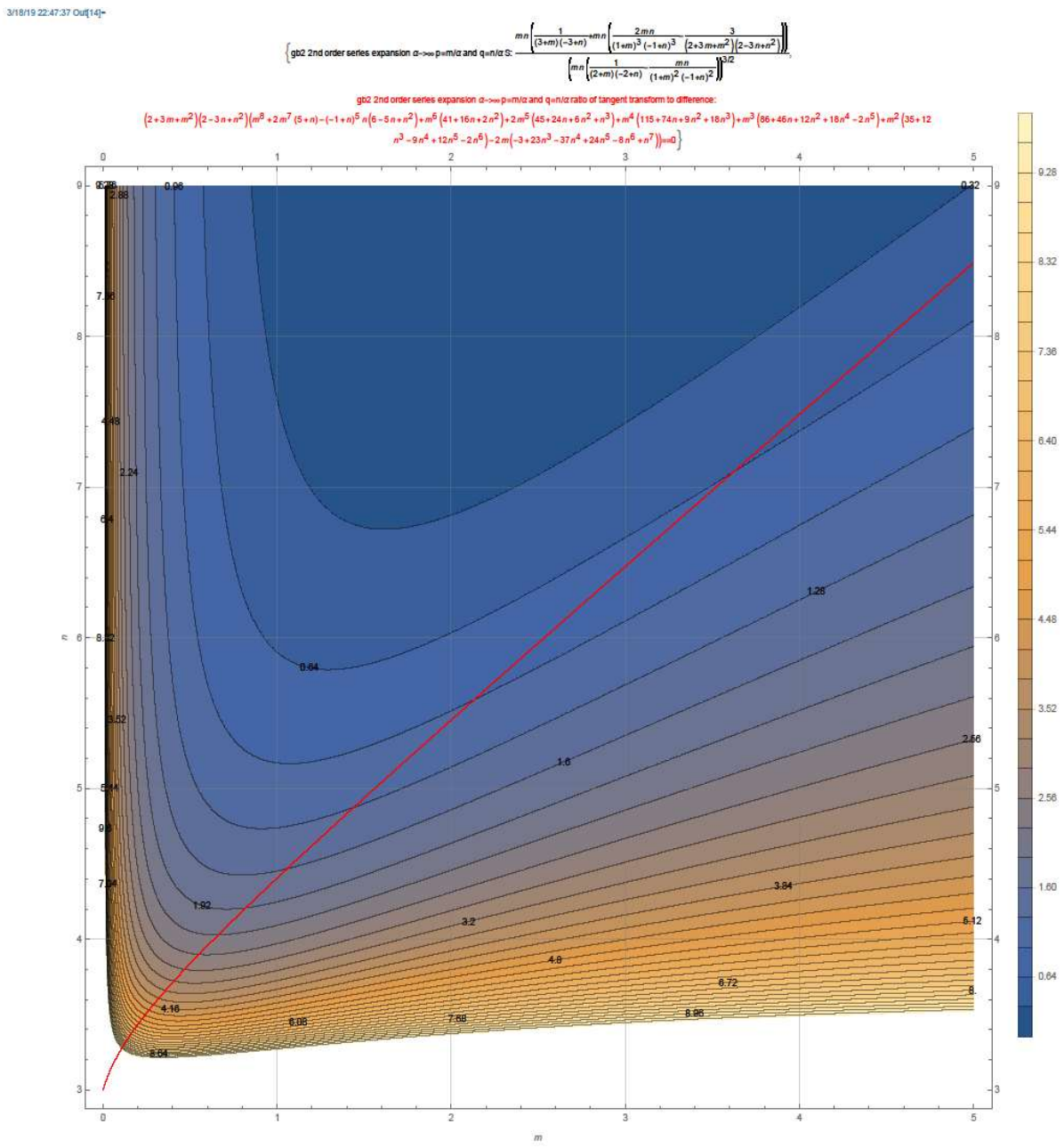
$$n > 4 \;\&\&\; m == Root[6 - 23n + 34n^2 - 24n^3 + 8n^4 - n^5 + (23 + 6n - 10n^2 + 6n^3 - n^4)\#1 + (34 + 10n)\#1^2 + (24 + 6n)\#1^3 + (8 + n)\#1^4 + \#1^5 \;\&, 3] \quad ...(37)$$

The S contour and the maximum *SF* location curve by Equation (35) plot is Figure 20.

The maximum *K* given *S* as calculated by Equation (36), and the McDonald Pade-approximation of the GB2 upper bound (McDonald *et al.* (2011) appendix), are in Figure 21.

When *n* increase from 4 to infinity, Equations (32-34) and (37) say that *S* decrease from 2.303367584519916 to zero, *K* decrease from infinity to 6, *SF* decrease from infinity to the minimum value 9.8414 when $n = 4.79175$ and $m = 1.3619$, with $S = 1.33373$, $K = 17.5064$, and then increase to infinity. The maximum *S* is 2.303367584519916 (McDonald *et al.* (2011) A.2 put a value 2.3037, a little less accurate). The minimum *K* is 6, which can be deduced analytically from the limit of the solution m in Equation (37) divided by n converges to one when n turns to infinity.

Figure 20: GB2 S Contours and Maximum SF Location Curve Plot When α → ∞ and p = m/α, q = n/α



From Figure 21 and similar plots zooming to different *S* ranges, we see that our maximum *K* is higher than McDonald *et al.* (2011)'s empirical formula when *S*<0.25, lower than it when 0.25<*S*<1, indistinguishable from it when 1<*S*<2.2, and lower than it when 2.2<*S*<2.3. The Pade approximation of McDonald *et al.* (2011) is 0 when *S* = 0.0600673 and is ∞ when *S* = 0.0600951, which may be the reason of its singular behavior near *S* = 0 and the wrong maximum *K* value less than 6.

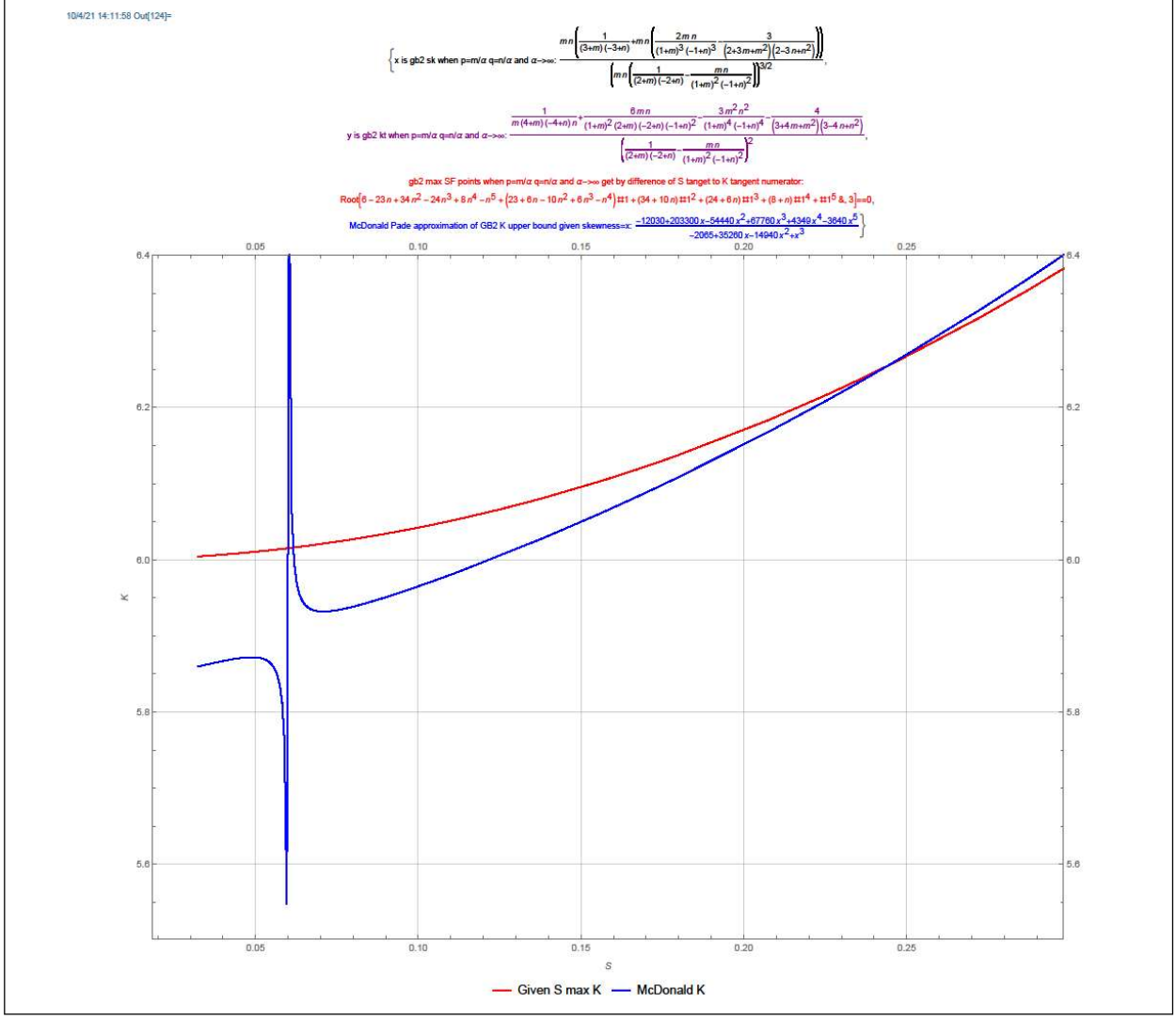
2.4.3. PDF Limit

Take symbolic limit individually for different parameter sub-intervals, or after taking logarithm and derivative, we get:

$$\lim_{\alpha\to\infty}\left(1+x^{\alpha}\right)^{\frac{1}{\alpha}}=\begin{cases}1 & 0\le x\le 1\\ x & x>1\end{cases} \qquad ...(38)$$

The PDF of $BetaPrimeDistribution[p,\ q,\ a,\ 1]$ is $\frac{\alpha\left(1+x^{\alpha}\right)^{-p-q}x^{-1+p\alpha}}{Beta[p,q]}$. Use Equation (38) and

Figure 21: GB2 Maximum K Given S Plot (in Red) When $\alpha \to \infty$ and p = m/α, q = n/α



$Limit\left[\dfrac{\alpha}{Beta\left[\dfrac{m}{\alpha},\dfrac{n}{\alpha}\right]},\ \alpha\to\infty,\ Direction\to FromBelow,\ Assumptions\to m>0\ \&\&\ n>4\right]=\dfrac{mn}{m+n}$, we find that when $\alpha\to\infty$ and $p=m/\alpha, q=n/\alpha$, the limit of GB2 distribution PDF is:

$$\underset{\alpha\to\infty,\ p=\frac{m}{\alpha},\ q=\frac{n}{\alpha}}{limit} GB2PDF=\begin{cases}\dfrac{mn}{m+n}x^{-1+m} & 0\le x\le 1\\ \dfrac{mn}{m+n}x^{-1-n} & x>1\end{cases} \qquad ...(39)$$

PDF with Equation (39) form is called Double Pareto distribution (DP) in Reed (2001), which is a very active distribution in contemporary researches. Reed (2003), Reed and Jorgensen (2004), Mitzenmacher (2004), Ribeiro *et al.* (2010), Faris (2011), Okamota (2013), Shin *et al.* (2015), Toda (2017), Pfitzinger *et al.* (2018), and Shriram (2018) are some of the references for DP with topics ranging from theory to multitude applications. The skewness and kurtosis of DP when m>0 and n>0 are in Equation (40-41).

$$DPS=\frac{2\left((1+m)^3(2+m)+(1+m)^2\left(-7+m(3+m)\right)n+3\left(3+m^2\right)n^2+5(-1+m)n^3-(-1+m)n^4\right)}{(3+m)\left(m(2+m)+(-1+n)^2\right)^{3/2}(-3+n)\sqrt{\dfrac{mn}{(2+m)(-2+n)}}Sign[-1+n]} \qquad ...(40)$$

$$DPK = 3(2+m)(-2+n)(2(-3+n)(-2+n)(-1+n)^4 - m(-2+n)(-1+n)^3(-29+(-4+n)n) + 9m^5(2+n+3n^2) + m^6(2+n+3n^2) + 2m^4(32+n^2(48+(-3+n)n)) + 3m^2(-1+n)^2(38+n(-1+n(17+(-7+n)n))) + 2m^3(58+n(-51+n(63+n(-1+3n))))))/(m(3+m)(4+m)(m(2+m)+(-1+n)^2)^2(-4+n)(-3+n)n) \quad ...(41)$$

Using Mathematica Reduce function, we know that Equation 40 and 41 is identical to Equation 32 and 33. In other words, GB2 upper bound distribution is asymptotic equivalent to *DP* distribution, i.e., when $\alpha \to \infty$, *BetaPrimeDistribution*[*m*/α, *n*/α, α,1] ~ *DoubleParetoDistribution*[*m*, *n*].

Due to piecewise PDF, Mathematica cannot directly calculate moment of *DP*. But it can calculate the central moment of DP, for example, the variance, or CM[2] of *DP* is $\frac{mn\left(1+2m+m^2-2n+n^2\right)}{(1+m)^2(2+m)(-2+n)(-1+n)^2}$. We can then calculate the moment of DP from central moment. The M[1] to M[4] of *DP* is in Equation 42.

$$\{DPM[1], DPM[2], DPM[3], DPM[4]\} = \left\{\frac{mn}{(1+m)(-1+n)}, \frac{m^2n^2}{(1+m)^2(-1+n)^2} + \frac{mn\left(1+2m+m^2-2n+n^2\right)}{(1+m)^2(2+m)(-2+n)(-1+n)^2}, \frac{m^3n^3}{(1+m)^3(-1+n)^3} + \frac{3m^2n^2\left(1+2m+m^2-2n+n^2\right)}{(1+m)^3(2+m)(-2+n)(-1+n)^3} + \frac{2mn\left(2+7m+9m^2+5m^3+m^4-7n-11mn+5m^3n+m^4n+9n^2+3m^2n^2-5n^3+5mn^3+n^4-mn^4\right)}{(1+m)^3(2+m)(3+m)(-3+n)(-2+n)(-1+n)^3}, \frac{m^4n^4}{(1+m)^4(-1+n)^4} + \frac{6m^3n^3\left(1+2m+m^2-2n+n^2\right)}{(1+m)^4(2+m)(-2+n)(-1+n)^4} + \frac{8m^2n^2\left(2+7m+9m^2+5m^3+m^4-7n-11mn+5m^3n+m^4n+9n^2+3m^2n^2-5n^3+5mn^3+n^4-mn^4\right)}{(1+m)^4(2+m)(3+m)(-3+n)(-2+n)(-1+n)^4} + \frac{3mn\left(12+58m+114m^2+116m^3+64m^4+18m^5+2m^6-58n-195mn-231m^2n-102m^3n+9m^4n+m^5n+114n^2+231mn^2+171m^2n^2+126m^3n^2+96m^4n^2+27m^5n^2+3m^6n^2-116n^3-102mn^3-126m^2n^3-2m^3n^3-6m^4n^3+64n^4+96m^2n^4+6m^3n^4+2m^4n^4-18n^5+9mn^5-27m^2n^5+2n^6-mn^6+3m^2n^6\right)}{(1+m)^5(2+m)(3+m)(4+m)(-4+n)(-3+n)(-2+n)(-1+n)^5}\right\} \quad ...(42)$$

In summary, the parameters transformation that is used to overcome the numerical optimization difficulties in distribution fitting Wang (2018b) reveals useful contours. The contour analysis is a form of asymptotic analysis in the transformed new parameters space. Different transformations or different contours show different properties of the distribution, some are for minimum of *SF*, such as Equation (31). Others are useful for locating maximum of *SF*, such as Equations (32-34), resembling our observation of the directional different characteristic numbers of GB2 distribution (Wang [2020] section 6.4). It is also of interest to notice that the useful contour form, such as $\gamma = \frac{k}{\alpha^a}$ or $p = \frac{m}{\alpha}, q = \frac{n}{\alpha}$, is in a format reminiscent of the shape factors construction $\frac{K}{|S|^\alpha}$.

2.5. GB1 Contour Analysis Part Two

2.5.1. When $\gamma \to \infty$ and $\alpha = m/\gamma$, $\beta = n/\gamma$

In section 2.1 and 2.2 we see GB1 when $\beta \to \infty$ has no difference with GB2. On the opposite, when $\beta \to 0$, GB1 is different from GB2 with $q \to 0$. The relevant contour analysis of GB2 expose its upper bound property to us in section 2.4. This prompt us to study the counterpart property of GB1 that when $\gamma \to \infty$, m>0, n>0, what is *GeneralizedBetaDistributionI*[*m*/γ, *n*/γ, γ, 1].

The GB1 PDF converges to a mixture of continuous and discrete distribution, in Equation (43) (using piecewise CDF construction or MixtureDistribution in Mathematica do not work, the direct expression utilizing DiracDelta function is the solution). The *S*, *K*, and *SF* is calculated in Equations (44-46).

$$ProbabilityDistribution\left[\frac{mn}{m+n}x^{-1+m} + \frac{m}{m+n}DiracDelta[x-1], \{x, 0, 1\}, Assumptions \to m > 0 \,\&\&\, n > 0\right] \quad ...(43)$$

$$S = -\frac{2\left(3m^3+3m^2(2+n)-n(3+n)+m\left(3+n^2\right)\right)}{(3+m)\sqrt{\frac{mn}{2+m}}(2+2m+n)^{3/2}} \quad ...(44)$$

$$K=\frac{3(2+m)\left(8m^5+2n^2(4+n)+8m^4(3+2n)-mn\left(8-4n+n^2\right)+12m^3\left(2+2n+n^2\right)+m^2\left(8+8n^2+3n^3\right)\right)}{m(3+m)(4+m)n(2+2m+n)^2} \quad ...(45)$$

$$SF=\frac{3(2+m)\left(8m^5+2n^2(4+n)+8m^4(3+2n)-mn\left(8-4n+n^2\right)+12m^3\left(2+2n+n^2\right)+m^2\left(8+8n^2+3n^3\right)\right)}{4(4+m)\left(3m^3+3m^2(2+n)-n(3+n)+m\left(3+n^2\right)\right)^2} \quad ...(46)$$

The difference of S contour tangent to the K contour tangent, $\frac{n(m+n)^2(2+2m+n)}{m(3+m)(4+m)\left(2m^2-2n+m(2+n)\right)}$, cannot equal zero. The minimum and maximum SF given S can only be at the $m-n$ parameter space boundaries.

At the upper boundary where $n\to\infty$, S is $-\frac{2(-1+m)}{\sqrt{\frac{m}{2+m}}(3+m)}$ and K is $\frac{3(2+m)\left(2-m+3m^2\right)}{m(3+m)(4+m)}$, which looks similar to *ParetoDistribution*[k, m, μ] S: $\frac{2\sqrt{\frac{-2+m}{m}}(1+m)}{-3+m}$ and K: $\frac{3(-2+m)\left(2+m+3m^2\right)}{(-4+m)(-3+m)m}$. But this is in reality the *PF* or *BetaDistribution*[m, 1]. When S>0, the K vs S plot is lower than *LogNormalDistribution*[μ, σ] plot, Figure 22. The corresponding maximum SF vs S plot is in Figure 23.

Figure 22: GB1 Maximum K Given S Plot (in Red) When $\gamma\to\infty$ and $\alpha=m/\gamma$, $\beta=n/\gamma$

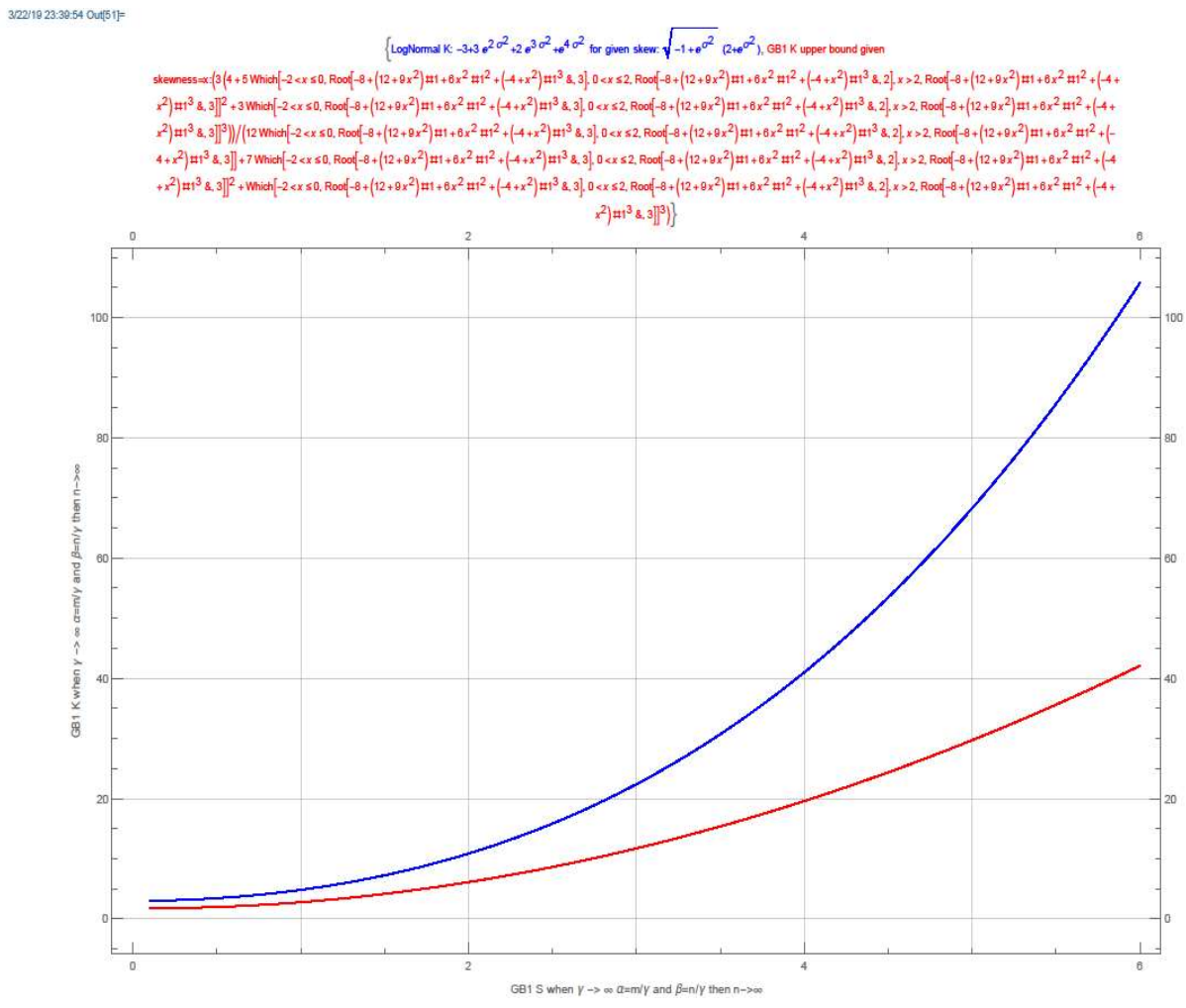


Figure 23: GB1 Maximum SF Given S Plot (in Red) When $\gamma \to \infty$ and $\alpha = m/\gamma$, $\beta = n/\gamma$

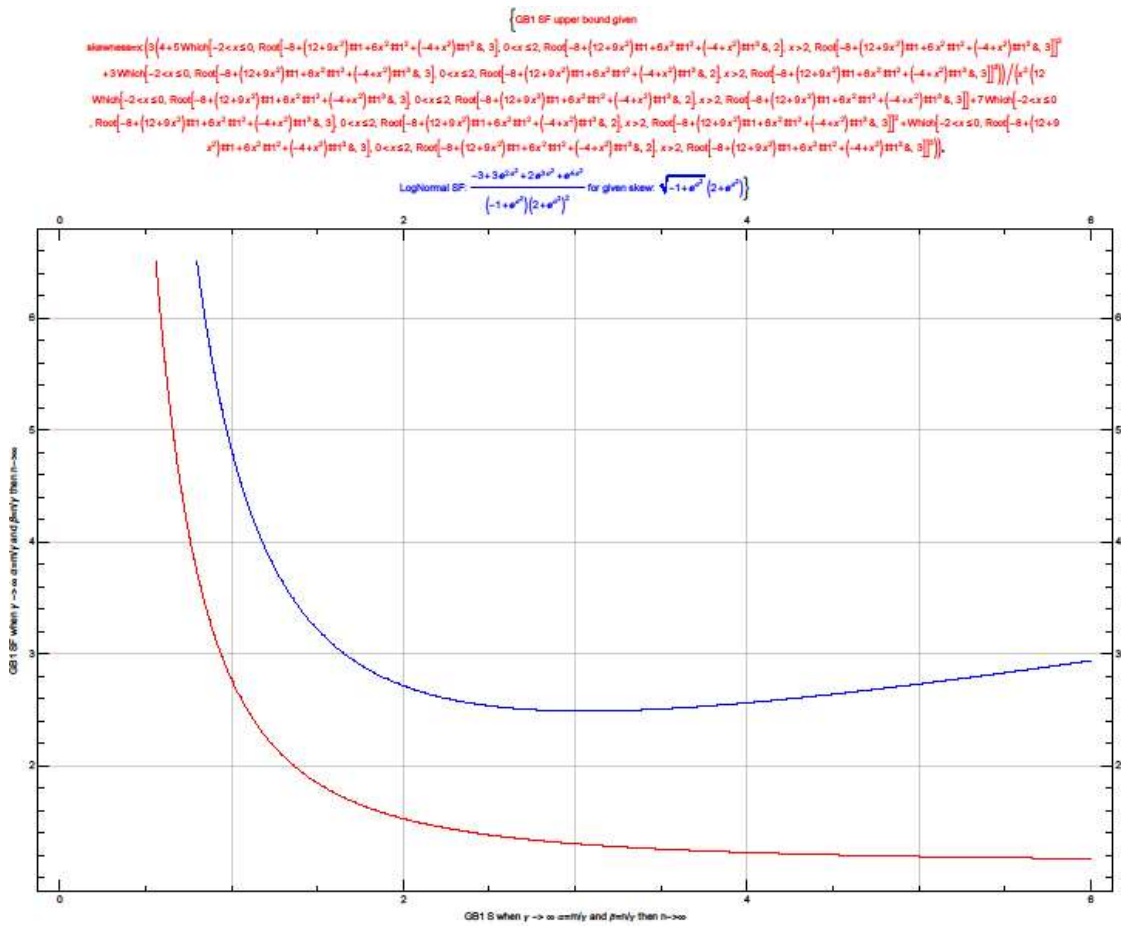


Figure 24: GB1 When $\gamma \to \infty$ and $\alpha = m/\gamma$, $\beta = n/\gamma$ Minimum SF Given m Plot, in Red if m<0.2, in Blue if m>0.2

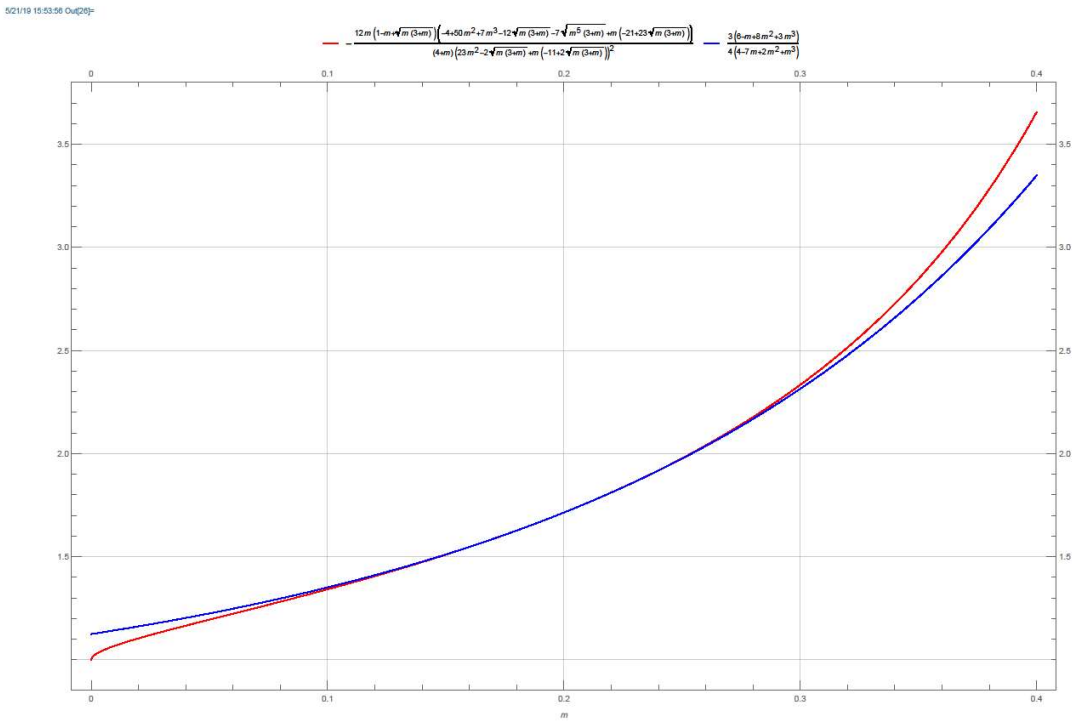
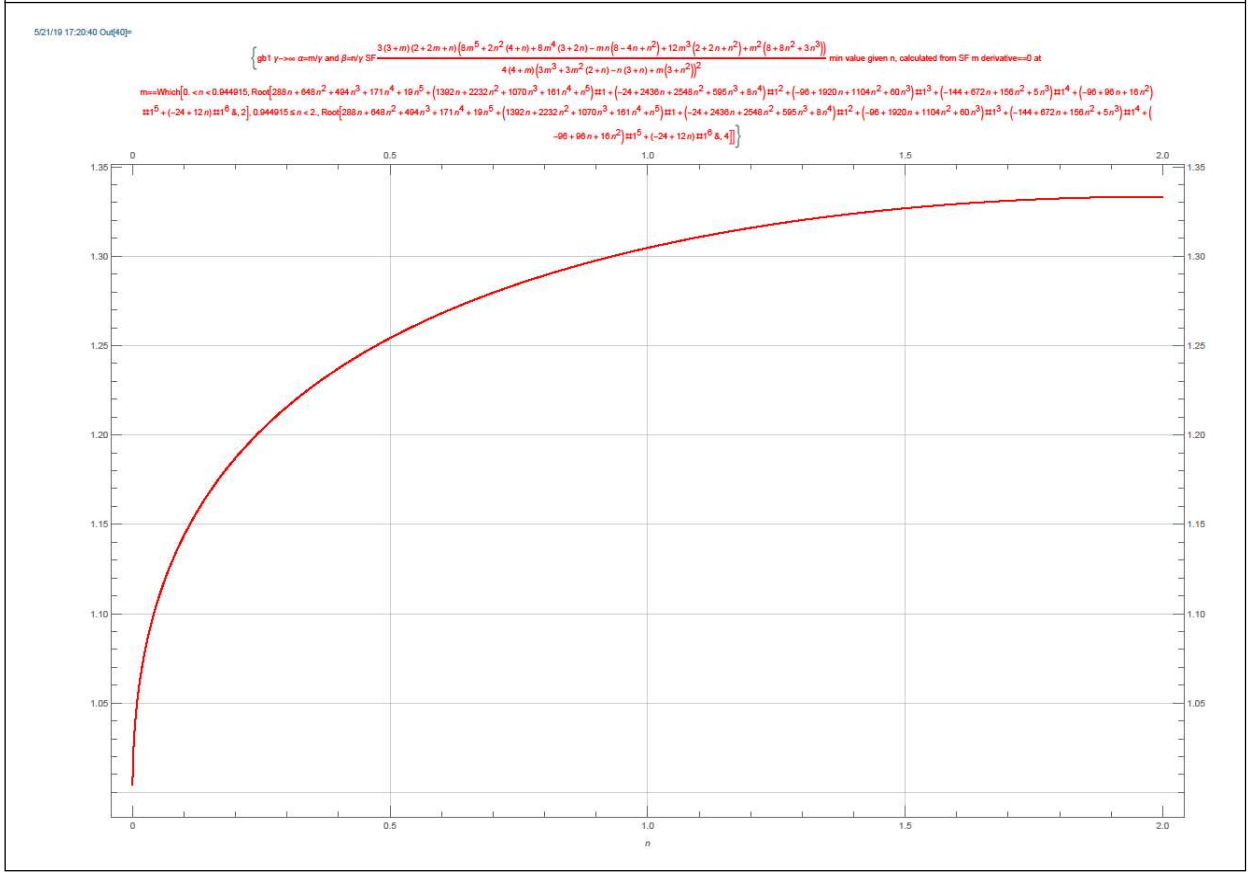


Figure 25: GB1 When γ → ∞ and α = m/γ, β = n/γ Minimum SF Given n Plot if n<2



Unlike the GB2 case, this GB1 counterpart contour maximum is dwarfed by the contour $\beta \to \infty$, $\alpha \to \infty$ and $\gamma \to 0$ results in section 2.2. One possible use of this limit is when finite range distribution with simple form are desired that are not as inclusive as GB1 but nonetheless extensive than Beta distribution. This mixture of continuous and discrete distribution bears resemblance to reinsurance loss distribution that has atom at the limit loss.

The minimum *SF* given m is attained at interior $n = -\frac{8\left(m+m^2\right)}{-1+5m} + 2\sqrt{\frac{3m+7m^2+5m^3+m^4}{\left(-1+5m\right)^2}}$ when 0<m<0.2, and at the upper boundary when m>0.2, through the *S* and *K* contour plot or the difference of *S* tangent to *K* tangent contour plot. The plot is in Figure 24, with $\beta \to 0$, which is simpler than Figure 12 from our section 2.2 contour analysis with $\beta \to \infty$, the antithesis asymptotic limit. When given n between 0 and 2, the minimum *SF* is also attained at interior m values, Figure 25, but it is at the negative *S* region and thus of less practical importance.

2.6. GB1 and GB2 Shape Factor Bound Plots Altogether

By now, we finished our GB1 and GB2 shape factor bound study via our asymptotic analysis. Our shape factor concept at the first thought may be regarded as merely combining two statistics of skewness and kurtosis into one characteristics. In reality, this simplification open the way to study the bound and range of shape factor. The boundary of the range, the minimum and the maximum, defined the suitable territory of each distribution family. Additionally, the asymptotic shape factor values for given parameter reveal the intrinsic meaning of the parameter. It tell us that different distribution families may possess similar fitting capabilities, the distinctive forms of the distributions may not be as important as we generally perceived. The parameter values uniquely determined per each asymptotic equivalent distribution class have meaning that is more essential.

With our exact analytical results of GB1 and GB2 shape factor bound, and of other comparison distributions bounds (to be published later) obtained using our presented methodology, we plot the shape factors bounds curves in Figure 26 and in Figure 27 with log scale. Our plots are amelioration of the empirical plot of Figures 1 and 2 in McDonald *et al.* (2011) that are highly consulted for distribution selection.

Figure 26: GB1 GB2 and Other Distributions Shape Factors Lower and Upper Bound Plot

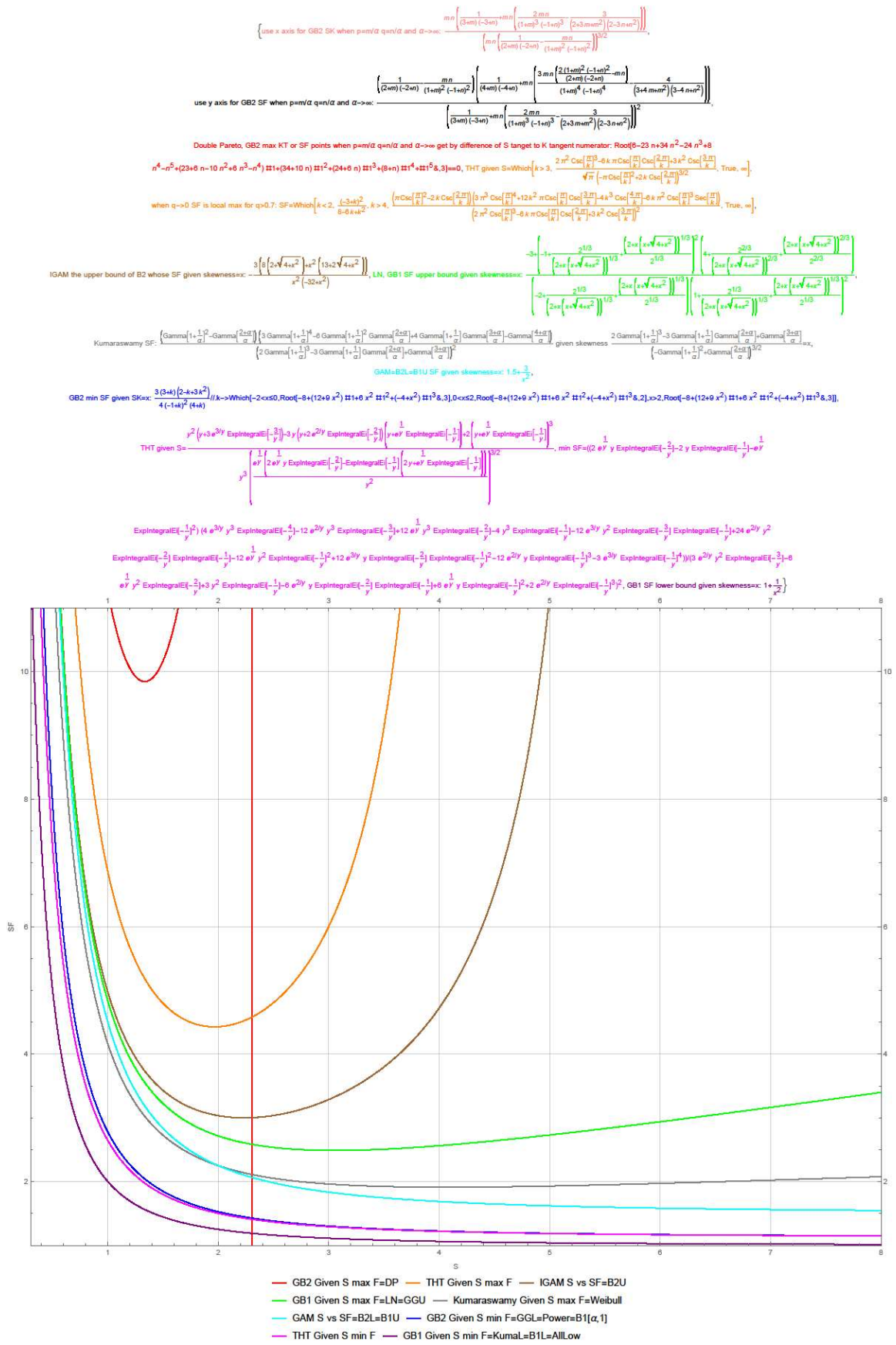
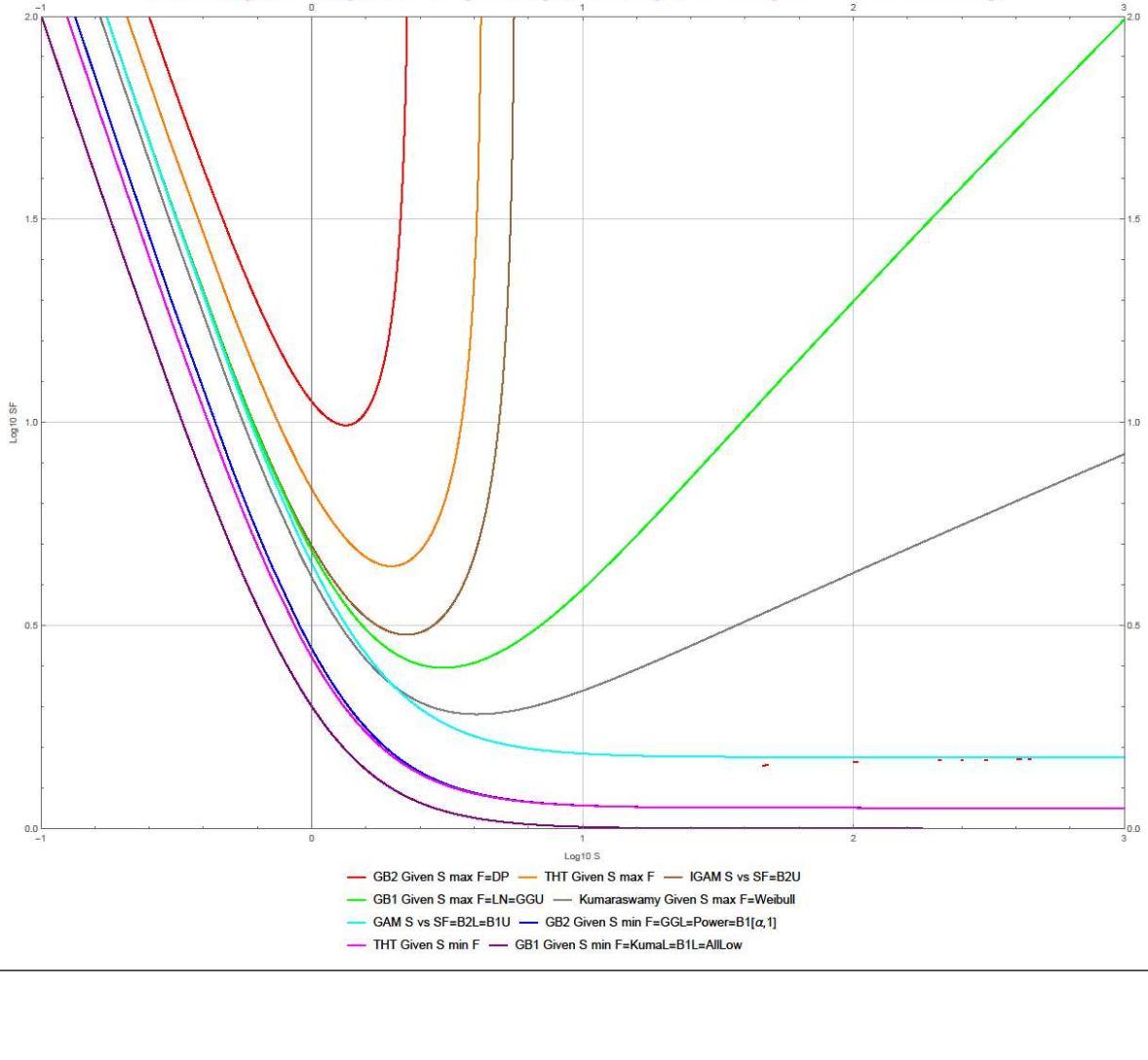


Figure 27: GB1 GB2 and Other Distributions Shape Factors Lower and Upper Bound in Log Scale



The precise value in our Figure 26 enable us to pinpoint the most suitable distribution in practice. As an example, for the aviation arrival-delay time distribution, which has a shape factor value of 2.016 and skewness 6.68, the best distribution is Weibull and THT. The latter turned out better captured the empirical distribution PDF shape near the lower endpoint.

3. Discussions and Conclusion

While the global minimum of the shape factor can be used to filter out not suitable probability distribution families, the conditional minimum or maximum of the shape factor (as well as the skewness and kurtosis) can locate the appropriate parameter ranges for distribution fitting. The usually simpler asymptotic limit formula may be employed to find such conditional minimum or maximum when it is attainted at the parameter space boundary. The asymptotic limit need to be replaced by contour limit in case the minimum or maximum is attainted at the parameter space corner.

The asymptotic or contour limit in simple cases may be obtained directly from taking symbolic limit in computer algebra system. In other cases, we can either use series expansion at various orders to simplify the expression first, or remove the subordinate terms in a sum expression, or expand the ratios of expressions by power series so that the other than power order factors can cancel out. In case there is no dominant term in a sum expression, and the series expansion of the terms in the sum expression annihilates each other and arrives at an infinitesimal result, we can try taking various logarithms and ratios to get the power or exponential order.

Whereas these symbolic limit tricks can get us results, they may also get us conflicting results, with contour limit at the corner more prone to error than the asymptotic limit at the boundary. One remedy for this is using semi-numeric method: combine the induction and deduction approach together by fixing some or all parameters to a list of numbers, and study the limit patterns with respect to the remaining parameters. If the special value substituted expression limits can be derived by symbolic limit now that fewer or no arbitrary symbol are involved, then we may be able to figure out the general limit formula forms. However, similar to a purely symbolic limit method may not work for expression involving general parameters, a purely numeric method may not work either: even reasonably large argument cannot be calculated when Gamma function is involved. Our formula Equations (29) and (30) are example application of the combined induction and deduction method.

When the conditional minimum or maximum of the shape factor is attained at the interior point, we can alternatively use the partial derivative contour plot or zero contour plot to locate the parameter positions, and calculate the corresponding shape factor values. The geometric or graphical analysis helps with symbolical analysis.

Once all these operational difficulties are successfully handled, the asymptotic and contour analysis techniques are not only helpful for cross validation of numerical plots or empirical formulas, but also capable of discovering new relationships or formulas, by providing information of where the shape factor minimum or maximum is taken place. That the upper bound of GB2 is attained by Double Pareto distribution is such an example. We believe other distributions are also amenable to these analyses.

After finished this research, in DP reference searching, we find that Higbee *et al.* (2019) already find that the DP is GB2 limit, in slightly generalized notation of their ALL. Other than our general framework of asymptotic analysis and induction with deduction techniques, our study contributed additional details as well as confirming their conclusion, such as where these maximum take place, Equations (35-37). Our approach may work even if the PDF limit do not exist. Besides, Higbee *et al.* (2019) do not find or mention that DP along the curve of Equation (35) is the GB2 upper bound, which is unidentified in McDonald *et al.* (2011).

To recapitulate from another perspective, the essence of our asymptotic analysis is complexity reduction: we cannot visualize three, four, or higher dimensional shape factor topography, but we can for one or two-dimensional parameters. Fix one or two parameters or expressions such as skewness or other meaningful form, in the GB2 and GB1 case the product of parameters and their powers, maximize or minimize shape factor over the remaining variables achieves this dimension reduction naturally. The simplification effect of asymptotic or contour analysis is accomplished either by taking boundary value or by taking specifically sliced sub manifold or contour limit when the minimum or maximum does not attain at the interior, such as Equations (31-34) or Figures 3 and 22, or alternatively, by locating partial derivative zero points when at the interior, such as Figures 4, 13, 20, 24, and 25.

In addition to this simplification, we believe boundary values are important because they in some sense determine interior values, as exemplified by harmonic or analytical function theories.

In signal processing, there is another definition of shape factor as the root mean square divided by the mean of the absolute value, https://www.mathworks.com/help/predmaint/ug/signal-features.html. This is a variant of CV, or the square root of the SF4[2] in Wang (2019a). We studied CV in Wang (2018a) and further research lead to our higher order shape factor definition that are more intrinsic to the shape of the distribution PDF, Wang (2018b and 2019a). The CV, Gini index, and normalized skewness, defined as the sign-keeping square root of the reciprocal of the shape factor, are indeed related to each other from our empirical study, which showed that CV is almost identical to Gini index, and are the base part or lower envelope of the normalized skewness; these are topics for a different research thesis.

Acknowledgment

We would like to thank all our friends who have encouraged and helped us in this study.

References

Faris M A (2011), Parameter Estimation for the Double Pareto Distribution. *Journal of Mathematics and Statistics,* 7(4): pp. 289-294.

Higbee J D, Jensen J E and McDonald J B (2019), The asymmetric log-Laplace distribution as a limiting case of the generalized beta distribution. *Statistics & Probability Letters,* 151: pp. 73-78. Available from: https://doi.org/10.1016/j.spl.2019.03.018

Marichev O and Trott M (2013), The Ultimate Univariate Probability Distribution Explorer. Available from: http://blog.wolfram.com/2013/02/01/the-ultimate-univariate-probability-distribution-explorer/

McDonald J B (1984), Some Generalized Functions for the Size Distribution of Income. *Econometrica,* 52(3): pp. 647-663.

McDonald J B, Sorensen J and Turley P A (2011), Skewness and kurtosis properties of income distribution models. LIS Working Paper Series, No. 569. Review of Income and Wealth. dx.doi.org/10.1111/j.1475-4991.2011.00478.x. Available from: https://pdfs.semanticscholar.org/eabd/0599193022dfc65ca00f28c8a071e43edc32.pdf

Mitzenmacher M (2004), Dynamic models for file sizes and double Pareto distributions. *Internet Mathematics,* 1(3): pp. 305-333.

Okamoto M (2013), Extension of the k-generalized distribution: new four-parameter models for the size distribution of income and consumption. Available from: https://pdfs.semanticscholar.org/a5e3/96cfb8d3da9d9c56b0bb6e62ccd67e65c907.pdf

Pfitzinger B, Baumann T, Emde A, Macos D and Jestädt T (2018), Modeling the GPRS network latency with a double Pareto-lognormal or a generalized Beta distribution. In: *Proceedings of the 51st Hawaii International Conference on System Sciences (HICSS).* doi: 0.24251/hicss.2018.730 Available from: http://hdl.handle.net/10125/50619

Reed W J (2001), The Pareto, Zipf and other power laws. *Econ. Lett*., 74: pp. 15-19.

Reed W J (2003), The Pareto Law of Incomes—An Explanation and an Extension. *Physica A,* 319: pp. 469-485.

Reed W J and Jorgensen M (2004), The double Pareto-lognormal distribution: A new parametric model for size distributions. *Communications in Statistics: Theory and Methods,* 33: pp. 1733-1753.

Ribeiro B, Gauvin W, Liu B and Towsley D (2010), On MySpace account spans and double Pareto-like distribution of friends. *Proc. IEEE INFOCOM* 2010(Mar.): pp. 1-6.

Shin W Y, Singh B C, Cho J and Everett A M (2015), A new understanding of friendships in space: Complex networks meet Twitter. *Journal of Information Science,* 41(6): pp. 751-764. Available from: https://doi.org/10.1177/0165551515600136

Shriram C K, Muthukumaran K and Murthy N L B (2018), Empirical Study on the Distribution of Bugs in Software Systems. *International Journal of Software Engineering and Knowledge Engineering,* 28(01): pp. 97-122. Available from: https://doi.org/10.1142/S0218194018500055

Toda A A (2017), A Note on the Size Distribution of Consumption: More Double Pareto than Lognormal. *Macroeconomic Dynamics,* 21: pp. 1508-1518. doi:10.1017/S1365100515000942

Wang F X (2018a), An inequality for reinsurance contract annual loss standard deviation and its application. In: Salman A and Razzaq M G A (eds.), *Accounting from a Cross-Cultural Perspective.* IntechOpen, 2018, pp. 73-89. doi: 10.5772/intechopen.76265. Available from: https://cdn.intechopen.com/pdfs/60781.pdf

Wang F X (2018b), What determine EP curve shape? Available from: http://dx.doi.org/10.13140/RG.2.2.30056.11523

Wang F X (2019a), What determines EP curve shape? In: Dr. Bruno Carpentieri (ed.) *Applied Mathematics.* dx.doi.org/10.5772/intechopen.82832. Available from: https://cdn.intechopen.com/pdfs/64962.pdf

Wang F X (2019b), Twisted Wang Transform Distribution. 2019 China International Conference on Insurance and Risk Management, Chengdu, China, July 17-20, 2019, Proceedings, pp. 912-926. doi:10.13140/RG.2.2.21901.49127. Available from: http://www.ccirm.org/conference/2019/CICIRM2019.pdf

Wang F X (2020), Shape Factor Asymptotic Analysis I. *Journal of Advanced Studies in Finance, Volume XI, Winter,* 2(22): pp. 108-125. doi:10.14505//jasf.v11.2(22).05 Available from: http://dx.doi.org/10.13140/RG.2.2.18427.36645

Cite this article as: Frank Xuyan Wang (2023). Shape Factor Asymptotic Analysis II. *International Journal of Pure and Applied Mathematics Research*, 3(1), 11-45. doi: 10.51483/IJPAMR.3.1.2023.11-45.