**Fast Bias-Removed Thêo1 Calculation with R**

W.J. Riley  
Hamilton Technical Services  
Beaufort, SC 29907 USA  
Rev. A June 28, 2020

- **Introduction**

This document describes a fast bias-removed Thêo1 calculation (ThêoBR) for the R statistical computing environment [1]. Some information about using R for frequency stability analysis will be found in Reference [2]. Thêo1 is an advanced statistic for analyzing frequency stability that provides results out to 75% of the phase data record length [3]. Some background about the evolution of the Thêo1, ThêoBR, and ThêoH statistics is given in Appendix 1. The ThêoBR calculation is implemented in C++ and is called from R using the Rcpp package [4]. It uses the fast Lewis recursive Thêo1 algorithm [5] along with the Taylor and Howe fast ThêoBR method [6]. Data sets of tens of thousands of points require only a few seconds rather than many minutes with the naive method. The calculation includes determination of one or two-sided error bars at a selected confidence factor. Tabulation and plotting of the results can be done within R, or an R driver function can be used for that purpose.

- **ThêoBR Function**

The ThêoBR function is defined as follows, and its C++ code is shown in Appendix 2.

```cpp
double TheoBR(NumericVector x, NumericVector t, NumericVector u,  
    NumericVector l, NumericVector tau, double tau0=1.0, double cf=0.683,  
    int br=1)
```

Arguments:
- `x` = NumericVector = Phase data input
- `t` = NumericVector = Theo1BR output
- `u` = NumericVector = Upper Theo1BR error bar
- `l` = NumericVector = Lower Theo1BR error bar
- `tau` = NumericVector = Tau output
- `tau0` = double = Data sampling time input (default=1)
- `cf` = double = Confidence factor  
  Negative = Single-Sided (default=0.683)
- `br` = integer = Bias remove flag (0=n0, 1=yes, default=1)

Return = double = Theo1 Bias Factor  
  or 1.0 if no bias correction  
  or 0 if error

The five NumericVectors are Rcpp data types comprise the input and output arrays, and they must be created in R before calling the TheoBR() function. The N point x array is the phase data, and the (N-1)/2 point t, u and l arrays are respectively the raw Thêo1 or ThêoBR, upper and lower error bar, and tau results. The `cf` parameter sets the confidence factor (negative for single-sided error bars) and the `br` flag determines whether the t output array is raw Thêo1 or bias-corrected ThêoBR. The phase data array is unchanged.
**ThêoBR Calculation**

The TheoBR() function (see Appendix 2) begins with an “all tau” Thêo1 calculation using the Lewis method (pasted directly from his Reference [5] paper except that the linear slope removal is moved inline), along with the corresponding tau values. While raw Thêo1 values are not required at every tau for the fast ThêoBR calculation, they are necessary for the fast recursive algorithm, and they can be made available without bias correction by setting the br argument to 0. The function then uses the Taylor and Howe fast ThêoBR method to determine the bias correction factor. That is applied to the raw Thêo1 values, and is also used to find the dominant integer power law noise type. Next, the noise type and number of data points is used to determine the edf, the number of equivalent chi-squared degrees of freedom that apply. Finally, the edf and confidence factor are used to calculate the applicable chi-square values and calculate the upper and lower error bars (the latter set equal to the nominal value for the single-ended case). The function ends by returning the correction factor.

**Overall Analysis**

The overall ThêoBR analysis is performed in R, and a typical procedure is as follows (R console commands are shown in red – beware of using the wrong quote character):

**To set up**

Install Rcpp with:

```r
> library(Rcpp)
```

Compile with:

```r
> sourceCpp("TheoBR.cpp")
```

after adding full path to filename e.g., "C:\R\Theo1BR.cpp"

**To use**

Load phase data into vector `x` (the tau0 must be known)

For example:

```r
> x<-scan("C:\Data\phase.dat")
```

Determine the # of phase data points with > N=length(x)

Allocate results vector `t` with:

```r
> t<-numeric((N-1)/2)
```

Allocate u vector tau with:

```r
> u<-numeric((N-1)/2)
```

Allocate l vector tau with:

```r
> l<-numeric((N-1)/2)
```

Allocate tau vector tau with:

```r
> tau<-numeric((N-1)/2)
```

**Call with**

```r
> Theo1BR(x, t, u, l, tau, tau0, cf, br)
```

where `tau0`, `cf` and `br` have defaults of 1.0, 0.683 and 1 respectively (either set their values or use the defaults)

**Display results**

```r
> t
```

or create a data frame

```r
> r<-data.frame(tau, l, t, u)
```
To tabulate the tau, minimum, nominal and maximum Theo1, and display it with:
> r

or plot the results with (for example, red line, no error bars):
> plot(tau,t,type="l",main="Theo1BR Plot",xlab="Tau",ylab="Theo1BR",
+ log="xy", col="red")

To add lines representing upper and lower error bounds:
> points(tau, u, type="l", col="blue")
> points(tau, l, type="l", col="blue")

To show the error bounds as a grey area on the plot:
> plot(tau,t,type="l",main="Theo1BR Plot",xlab="Tau",ylab="Theo1BR",
+ log="xy", col="red")
> polygon(c(tau, rev(tau)), c(u, rev(l)), col = "grey70", border = NA)
> points(tau,t,type="l",col="red")

Other R plot formats can be used as desired. An R driver function can encapsulate the complete analysis process (see Appendix 4).

• Examples

Note: These examples assume that the basic setup and array allocation has been done. If the results arrays are larger than required, the extra values should be ignored.

1. Check a basic Thêo1 calculation using test values from Reference [2].
   • Enter test values into x array.
     > ttest<-c(1.00,2.50,0.65,-3.71,-3.30,1.08,0.50,2.20,4.68,3.29)
   • Call ThêoBR() with br=0.
     > TheoBR(ttest,t,u,l,tau,1.0,0.683,0)
   • Examine the results in t array. The four raw Thêo1 results are for even averaging factors m from 2 to 8. The latter agrees with the m=8 value in the paper.
     > t
     [1] 2.055700408 1.509405466 1.412349249 1.148758425

2. Check the results of a ThêoBR() calculation using the 1000-point test suite of Reference [6] after conversion to 1001 points of phase data as compared with Stable32.

• Conclusions

The naïve Thêo1 algorithm that closely follows its definition is fine for determinations at a single tau or over a tau range for data set sizes up to 10,000 points or so. For example, an “All Tau” Thêo1 analysis for that number of phase data points takes about 20 seconds. But for 100,000 data points, the calculation takes about 30 minutes. Similar times apply to ThêoBR and ThêoH. The new Lewis fast Thêo1 algorithm changes that, making Thêo1 runs practical for very large data sets, as demonstrated here. The R/RStudio statistical computing environment combined with C++ via Rtools/Rcpp allows the Lewis Thêo1 algorithm to be implemented in a convenient way.
• References

4. The easiest way to use C/C++ code in R is with the Rcpp and Rtools tool chain. Rcpp supports calling C++ from R, while Rtools compiles C++ code under R. Rcpp.
5. B. Lewis, "Fast Algorithm for Calculation of Thêo1," Submitted to IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, May 2020. This paper presents a fast, recursive algorithm for “all tau” Thêo1 calculations.
Appendix 1

Notes for Thêo1, ThêoBR, and ThêoH

The Thêo1 statistic was formally introduced in the May 2003 Howe and Peppler paper of Reference [3], and by that time had been the subject of “discussions, interpretations and tests”, including implementation into the Stable32 program. That code was created and running in early February 2003 and was revised to the paper’s “final” version by the end of April of that year. The 2003 Howe and Peppler paper described Thêo1 as applying for even averaging factors, m, from 10 to N-1 (where N is the number of phase data points – the Lewis fast Thêo1 algorithm covers m from 2). It discussed its stride of $0.75m\tau_0$, where $\tau_0$ is the data sampling interval (and thus provides results out to 75% of the record length), compared its frequency response to that of the Allan variance, AVAR, gave simple estimates for its bias versus AVAR for various power law noise processes which ranged from factors between 0.77 to 1.50 for ADEV, and gave empirical formulae for its equivalent number of Chi-squared degrees of freedom (edf). Reference [9] followed with more details in 2004. Note that the Thêo1 name generally applies to the variance, while most usage is for its square root, the Thêo1 deviation, and it is important to keep that distinction in mind when discussing it.

Thêo1 is a biased estimator of the Allan variance for other than white FM noise, and, if the noise type is known, one way to determine the necessary correction factor is to use an empirical formula. References [3] and [7] included a preliminary version of that, and a later improved version is given in Reference [10], which also supported Thêo1 calculations down to m=2. The latter appears in the Stable32 as the Theo1Bias() function, and that Thêo1 bias expression also shown on page 30 of the Handbook of Frequency Stability Analysis [10].

Work continued on Thêo1, particularly regarding ways to determine its bias without knowledge of the power law noise type (or mixtures). The best approach was to simply calculate the average ratio of Thêo1 to AVAR over a range of averaging factors. That is the method used for bias-removed ThêoBR herein. The final approach for that is described in Reference [6].

However it is still necessary to know the dominant power law noise type in order to determine the edf and to set Chi-squared confidence limits for a Thêo1/BR result. The Thêo1/BR bias factor can be used to determine the noise type, and a function for that purpose was incorporated into Stable32 in December 2003 based on the T.N. Tasset & D.A. Howe paper, “A Practical Thêo1 Algorithm”. Then the expressions in Reference [2] can be used to determine the edf, and the desired single or double-sided error bars set accordingly. Thêo1 and ThêoH actually have narrower confidence limits as discussed in References [11] and [12].

The ThêoBR calculation in Stable32 does phase averaging to speed up the bias factor determination. That process was the subject of several e-mails between D.A. Howe and T.N. Tassset and the author between May 2003 and July 2004, and their (unpublished?) paper mentioned above. The July 2004 Stable32 code for automatically determining the ThêoBR bias correction includes the following notes:

```c
/* This bias factor is the ratio of AVAR to Thêo1 */
/* To correct the Thêo1 deviation, multiply it by the sqrt */
/* of this bias factor. */
/* */
/* Thêo1 is the original biased statistic of that name. */
/* NewThêo1 was the original bias-corrected version of */
/* Thêo1, and that bias correction was found with the */
/* AutoTheo1Bias() function. This bias-corrected version */
```
The ThéoBR bias correction function was later changed to calculate the bias correction from the longest averaging factor downwards.

The ThéoH statistic (see Reference [8]) is a combination of the overlapping Allan and ThéoBR deviations whereby bias-removed Théo1 is used to extend an ADEV analysis to longer averaging factors. The combination is used instead of a purely ThéoBR analysis mainly because the ADEV computation is faster for the shorter averaging times where it applies. The advent of the fast Lewis Théo1 algorithm, which has a speed advantage factor on the order of N (even though its recursive nature requires an “all tau” calculation), offers the possibility of replacing ThéoH with a bias-removed Théo1 run like that described herein. In effect, the fast ThéoBR deviation becomes a better ADEV (at least for a compete analysis run).

The reason for this discussion is to point out that, as the family of Théo1 statistics evolved, slightly different versions were developed from the same core Théo1 definition that use varying averaging factor ranges, bias correction methods and other details. The fast Lewis algorithm produces (very nearly) the same Théo1 estimate but the subsequent ThéoBR bias corrections and ThéoH presentations may differ between implementations.

The R ThéoBR results are obtained essentially instantly for N=10³, and nearly so for N=10⁴, while the latter takes about 20 seconds for a Stable32 “All Tau” (really “Many Tau”) Théo1 run, thus demonstrating the benefit of the Lewis fast algorithm. An N=10⁵ run takes a reasonable ≈ 12 seconds for the R ThéoBR results, while it is a rather impractical ≈ 30 minutes for the naive method in Stable32.

This ThéoBR C++ code from Reference [5] does not use 128-bit integers to avoid potential numerical issues for large data sets (see that paper for details).

It is worth emphasizing that there is nothing inherently wrong about using a biased estimator of a statistic when it offers superior properties like wider applicable range or better confidence provided that the bias can be removed. Similarly, there is no disadvantage for a more complicated estimator as long as it can be implemented effectively (e.g., the fast Lewis algorithm for Théo1).

As an aside, the R/RStudio computing environment, while it may not be ideal for most routine frequency stability analysis, is, along with Rcpp, an attractive experimental environment for developing C++ analysis software where the interactive R user interface serves as an effective top-level means for editing, compiling, test and evaluation.
Appendix 2

C++ Code for ThêoBR with R

The following C++ code may be copied, pasted into a text editor, and saved as TheoBR.cpp in an appropriate folder, e.g., C:\R\ on a Windows® system, where it can be accessed from R.

```c++
/* TheorBR.cpp */
/* All Tau Bias Removed Theo1 Calculation for R */
/* Using Fast Lewis Theo1 and Taylor/Howe Fast TheoBR Methods */
/* */
/* This code calculates a bias-removed all tau version of Theo1 for R, */
/* based on the Lewis fast Theo1 algorithm and the Taylor/Howe fast Theo1 */
/* method. It is completely contained within this file. The N point phase */
/* data x[], and the (N-1)/2 point Theo1 t[], upper [u] and lower [l] error */
/* bar, and Theo1 tau[] array arguments are NumericVectors. The data */
/* sampling interval tau0 is input as a double with a default value of 1, */
/* the confidence factor as a double with default value 0.683, and the bias */
/* removal flag as an integer (0 or 1) with default value 1. The x[] is */
/* unchanged, t[] receives the Theo1BR results, u[] and l[] receive the */
/* upper and lower Theo1 error bar values, tau[] receives the corresponding */
/* Theo1 tau values, and the function returns a double, either the Theo1 */
/* bias factor, 0 if no correction, or an error code. The tau values are */
/* 0.75 * tau0 * even averaging factors. If the last br argument is 0, */
/* the function returns 1.0 with the raw uncorrected Theo1 values. Chi- */
/* squared error bars are calculated for each AF based on the # of data */
/* points, the Theo1 EDF, and the desired CF. A negative CF entry denotes */
/* use of single-sided confidence limits. */
/* */
/* To set up: */
/* Install Rcpp with: > library(Rcpp) */
/* Compile with: > sourceCpp(“Theo1BR.cpp”) */
/* after adding full path to filename e.g., ”C:\R\Theo1BR.cpp” */
/* */
/* To use: */
/* Load phase data into vector x (the tau0 must be known) */
/* e.g., > x<-scan(”C:\Data\phase.dat”) */
/* Determine the # of phase data points with > N=length(x) */
/* Allocate results vector t with: > t<-numeric((N-1)/2) */
/* Allocate u vector tau with: > u<-numeric((N-1)/2) */
/* Allocate l vector tau with: > l<-numeric((N-1)/2) */
/* Allocate tau vector tau with: > tau<-numeric((N-1)/2) */
/* */
/* Call with: */
/* > Theo1BR(x, t, u, l, tau, tau0, cf, br) */
/* where tau0, cf and br have defaults of 1.0, 0.683 and 1 respectively */
/* */
/* Display results with: */
/* > t or create a data frame */
/* > r<-data.frame(t,u,l,tau) and display it with > r */
```
/* or plot the results with (for example, red line, no error bars):        */
/* > plot(tau,t,type="l",main="Theo1BR Plot",xlab="Tau",ylab="Theo1BR",   */
/*   + log="xy", col="red")                                         */
/* To add lines representing upper and lower error bounds:              */
/* > points(tau, u, type="l", col="blue")                             */
/* > points(tau, l, type="l", col="blue")                             */
/* To show the error bounds as a grey area on the plot:                */
/* > plot(tau,t,type="l",main="Theo1BR Plot",xlab="Tau",ylab="Theo1BR",  */
/*   + log="xy", col="red")                                                   */
/* > polygon(c(tau, rev(tau)), c(u, rev(l)), col = "grey70", border = NA) */
/* > points(tau,t,type="l",col="red")                                   */
/* Other R plot formats can be used as desired                         */
/* An R wrapper function can encapsulate the complete analysis process */
/* References:                                                           */
/*      to IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency */
/*      Control, May 2020.                                               */
/*     Set Stability Analysis", IEEE Transactions on Ultrasonics, Ferro-  */
/*     electrics, and Frequency Control, September 2010.                */

/* Revision Record:                                                    */
/* 06/10/20  Created and running                                        */
/* 06/16/20  Added draft error bar code                                 */
/* 06/17/20  Editing and debugging                                      */
/* 06/18/20  Changed name to TheoBR()                                   */
/* 06/19/20  Added provisions for single-sided confidence limits        */
/* 06/23/20  Validated and documented                                   */

/* W.J. Riley, Hamilton Technical Services, Beaufort, SC 29907 USA © 2020 */
/* E-mail: bill@wriley.com                                             */
/* License: MIT                                                        */

/***************************************************************************/
#include <Rcpp.h>
using namespace Rcpp;

/***************************************************************************/
/* Calculate TheoBR for a phase dataset taken at time increments tau0. The */
/* phase data are copied to a working array X, and any linear component to */
/* the dataset is removed as a first step. The raw all tau Theo1 values are */
/* are calculated using the Lewis fast algorithm of Reference [1], the correspond corresponding tau values are written, and Theo1BR bias corrections are */
/* made per Reference [2]. Chi-squared upper and lower error bars are */
/* returned in the u and l arrays for the given confidence factor. The # of */
/* Theo1 results is (N-1)/2. This code should preferably be compiled with: */
/* -O3 -ffast-math.                                                        */
/* Arguments:                                                            */
/* x = NumericVector = Phase data input */

8
t = NumericVector = Theo1BR output
u = NumericVector = Upper Theo1BR error bar
l = NumericVector = Lower Theo1BR error bar
tau = NumericVector = Tau output
tau0 = double = Data sampling time input (default=1)
cf = double = Confidence factor
br = integer = Bias remove flag (0=no, 1=yes, default=1)

Return = double = Theo1 Bias Factor
          or 1.0 is no bias correction
          or 0 if error
****************************************************************************

// [[Rcpp::export]]

// The core Theo1BR computational routine was created by
// Ben Lewis of the University of Strathclyde UK
// b.lewis@strayh.ac.uk per Reference [1] above © 2020 IEEE

// TheoBR function
double TheoBR(NumericVector x, NumericVector t, NumericVector u,
    NumericVector l, NumericVector tau, double tau0=1.0, double cf=0.683,
    int br=1){

    // Find phase array size
    int N=x.size();

    // Copy phase data to working array
double* X = new double[N];
    for(int i = 0; i < N; i++){
        X[i] = x[i];
    }

    // Initializations
    int k_max = (N-1)/2;
    int single = 0; // Flag for single-sided confidence limits
    if(cf<0){
        single=1;
        cf = -cf;
    }
    double* C1 = new double[N];
    double* C3 = new double[k_max+1];
    double* C4 = new double[k_max*2];

    // Preprocess by removing linear part
    double midpoint = ((double) (N-1))/2;
    long double sum1 = 0;
    long double sum2 = 0;
    for(int i = 0; i < N; i++){
        sum1 += X[i];
        sum2 += X[i]*(i-midpoint);
    }

    // Calculations
    // ...

    // Return Bias Factor
    if(single){
        //...
    } else {
        //...
    }
}

double a = sum1/N;
double b = sum2/N*12/((double)N*N-1);
for(int i = 0; i < N; i++){
    X[i] = a + b*(i-midpoint);
}

// Calculate C1
double s=0;
for(int i = 0; i < N; i++){
    s += (X[i]*X[i]);
    C1[i] = s;
}

// Main loop
C3[0] = C1[N-1];
for(int k=1; k<=k_max; k++){

    //Calculate C2 values
    double C2_2k = 0;
    double C2_2k_1 = 0;
    for(int j = 0; j <= N-2*k-1; j++){
        C2_2k += (X[j]*X[j+2*k]);
        C2_2k_1 += (X[j]*X[j+2*k-1]);
    }
    C2_2k_1 += (X[N-2*k]*X[N-1]);

    // Update C3, C4 in place
    for(int v=0; v < k; v++){
        C3[v] -= (X[k-1-v]*X[k-1+v]) + (X[N-k+v]*X[N-k-v]);
    }
    for(int v = 1; v<2*k-2; v++){
        C4[v-1] -= (X[2*k-1-v]*X[2*k-1]) + (X[2*k-2-v]*X[2*k-2]) + (X[N-2*k]*X[N-2*k+v]) + (X[N-2*k+1]*X[N-2*k+1+v]);
    }
    C3[k] = C2_2k;
    C4[2*k-2] = 2*C2_2k_1 - (X[0]*X[2*k-1]) - (X[N-2*k]*X[N-1]);
    C4[2*k-1] = 2*C2_2k;

    //Calculate un-normalised T_k from C1-C4
    double T_k = 0;
    double A0 = C1[N-1] - C1[2*k-1] + C1[N-2*k-1] + 2*C2_2k;
    for(int v = 1; v <= k; v++){
        double A1 = A0 - C1[v-1] + C1[N-1-v] - C1[2*k-v-1] + C1[N-1-2*k+v];
        double A2 = C3[k-v] - C4[v-1] - C4[2*k-v-1];
        T_k += (A1+2*A2)/v;
    }
}
// For testing - check for negative Theo1
if(T_k<0) return 0;

// Apply normalisation to get Theo1
// t[k-1] = (T_k/(3*(double)(N-2*k)*k*k));
t[k-1] = (T_k/(3*(double)(N-2*k)*k*tau0*tau0));
if(br==0){
    // Take square root for Theo1 deviation
    t[k-1] = sqrt(t[k-1]);
}

// Enter tau = averaging factor * tau0
tau[k-1] = 0.75*2*k*tau0;

// Release memory
delete[] C1;
delete[] C3;
delete[] C4;

// Return raw Theo1 if no bias correction
if(br==0){
    return 1.0;
}

// Now have raw all tau Theo1 variances and their corresponding tau values.
// Next need to determine the bias correction factor kf using n=(0.1*N/3)-3
// AVAR/Theo1 variance ratio averages per Ref [2] (other n's have been used,
// e.g., (N/6)-3). The AVAR and Theo1 variance values used are for the same
// taus, with AFs of 9+3i and 12+4i respectively where i goes from 0 to n
// (sometimes n to 0 has been used instead). The Theo1 variance values come
// from the t[] array with the index above.
double kf = 0.0;
int n = ((0.1*N)/3)-3;
int m; // Averaging factor

// Correction factor summation loop
for(int i=0; i<=n; i++){
    // Calculate AVAR at AF=m=9+3i
    m=9+3*i;
double avar = 0.0;
    for(int j=0; j<N-2*m; j++){
        // avar += pow((X[j+2*m] - 2*X[j+m] + X[j]), 2);
        avar += ((X[j+2*m] - 2*X[j+m] + X[j])*((X[j+2*m] - 2*X[j+m] + X[j])));
    }
    // Apply scale factor
    double M=m;
    avar /= (2*M*M*(N-2*M)*tau0*tau0);
}

// For testing
// Divide AVAR by corresponding Theo1 variance
// and add that to correction factor sum
kf += (avar / t[5+2*i]);
}
// Divide kf sum by # ratios
kf /= (n+1);
// For testing
// Rprintf("kf=%e
", kf);
// For testing - Force kf=1.0
kf=1.0;
// Apply bias corrections and convert to Theo1 deviation.
for(int i=0; i < k_max; i++){
t[i] *= kf;
t[i] = sqrt(t[i]);
}

// To set error bars, must determine
// dominant power law noise type from
// Theo1 bias correction factor, then
// calculate Theo1 edf, and use that to
// determine upper and lower chi-squared
// confidence limits at (say) 68%
// 1-sigma Confidence Factor

// Use Theo1 bias factor to determine noise type
// There is one bias factor value and one noise
// type determination for the whole run
// Thresholds for noise sorting are geometric means
// of nominal Theo1 variance bias values

// Note: Nominal Theo1 variance bias factors in
// original 2003 FCS paper were:
// WPM=0.4, FPM=0.6, WFM=1.0, FFM=1.71, RWFM=2.24
// These are not used here

// Error bar determination loop
for(int i=0; i < k_max; i++){
    // Get nominal Theo1 var bias factors for various
    // power law noise types versus averaging factor
    m=i+1;
    // double mm = 0.75*m;
double mm = 0.75*2*m;
double bwPM = (0.09 + (0.74/pow(mm, 0.40)));
double bfPM = (0.14 + (0.82/pow(mm, 0.30)));
double bwFM = 1.00;
double bFFM = (1.87 + (-1.05/pow(mm, 0.79)));
double bRWFM = (2.70 + (-1.53/pow(mm, 0.85)));
}
// For testing
// Display the nominal bias factors
// Rprintf("mm=%d, WPM=%f, FPM=%f, WFM=%f, FFM=%f, RWFM=%f\n",
//   mm,bWPM,bFPM,bWFM,bFFM,bRWFM);

// Find noise type
int a; // Power law noise exponent alpha
if(kf<sqrt(bWPM*bFPM)) a=2;
else if(kf<sqrt(bFPM*bWFM)) a=1;
else if(kf<sqrt(bWFM*bFFM)) a=0;
else if(kf<sqrt(bFFM*bRWFM)) a=-1;
else a=-2;

// Find edf
// N is # phase data points
// mm is tau-s the Theo1 stride = 0.75*m
double edf; // Theo1 EDF
if(a==2) edf=(0.86*(N+1)*(N-((4.0/3.0)*mm)))/((N-mm)*(mm+1.14));
if(a==1) edf=((4.798*N*N)-(6.374*N*mm)+(12.387*mm))/
   (sqrt(mm+36.6)*(N-mm)*(mm/(mm+0.3)));
if(a==0) edf=(((4.1*N+0.8)/mm)-((3.1*N+6.5)/N))*
   (pow(mm,1.5)/(pow(mm,1.5)+5.2));
if(a==-1) edf=((2.0*N*N-1.3*N*mm-3.5*mm)/
   (N*mm)*((mm*mm*mm)/(mm*mm*mm+2.3)));
if(a==-2) edf=((4.4*N-2)/(2.9*mm)*(((4.4*N-1)*
   (4.4*N-1)-8.6*mm*(4.4*N-1)+
   11.4*mm*mm)/(4.4*N-3)/(4.4*N-3)));
edf=(((4.1*N+0.8)/mm)-((3.1*N+6.5)/N))*(pow(mm,1.5)/(pow(mm,1.5)+5.2));

// Find inverse upper and lower Chi-squared values for given confidence
// factor and edf with an embedded R calls.
// See: https://teuder.github.io/rcpp4everyone_en/
// 220_dpqr_functions.html#chi-squared distribution
// Set confidence interval and fill in upper and lower error bar arrays
// Lower error bar equals nominal TheoBR for single-sided conf interval
if(single){
    double upper = R::qchisq((cf),edf,0,0);
    u[i] = sqrt(t[i]*t[i]*edf/upper);
    l[i] = t[i];
}
else{ // Double-sided confidence intervals:
    double lower = R::qchisq(1-cf)/2,edf,0,0);
    double upper = R::qchisq(1-(1-cf)/2,edf,0,0);
    u[i] = sqrt(t[i]*t[i]*edf/upper);
    l[i] = sqrt(t[i]*t[i]*edf/upper);
}

// For testing
// if(i==4) Rprintf("tau=%f, min=%f, nom=%f, max=%f\n", tau[i],l[i],t[i],u[i]);
// Done
return kf;
}
Appendix 3

Validation of the ThêoBR() Function

1. Comparison of (raw, no bias correction, W FM) Thê1 deviation results at m=10, tau=7.5 for 12-point test suite of Reference [8] with that paper, Stable32, and this ThêoBR() function with br=0: All results agree (0.6623816).

2. Comparison of (raw, no bias correction, W FM) Thê1 deviation results for 1000-point test suite of Reference [13] converted to 1001 points of phase data between Stable32, and this ThêoBR() function with br=0: All results agree.

---

Note that these Stable32 Thê1 results begin at AF=10

Stable32 Results Thru tau=28.5

Note that these Stable32 Thê1 results are spaced for “Many Tau” rather than “All Tau”

---

ThêoBR() Results thru tau=28.5

---

Stable32 Results at largest taus

---

ThêoBR() Results at largest taus

15
3. Add a test code statement to show the ADEV values used for the ThêoBR bias factor determination: Rprintf("i=%d, m=%d, adev=%e\n", i,m,sqrt(avar)); The ADEV loop goes from i=0 to n where n = ((0.1*N)/3)-30 and N=# phase data points (1001). The AVAR values used for the average bias factor ratio are at m=9+3i (9 thru 99) and the Thêo1 variance are at m=12+4i (12 thru 132) which correspond to 0-bases t[] array indices 5+2i (5 thru 65). The Stable32 results are edited to remove AFs not applicable. Note that some ThêoBR AFS (m) are missing from the Stable32 “Many Tau” AFS. All results agree.

4. For this white FM noise test data, the ADEV and Thêo1 deviation values at the same tau should be approximately the same. At the minimum τ=9, they are 9.8404e-2 and 9.8141e-2 respectively, close to the same, and have an Allan/ Thêo1 variance ratio of 1.0054, very close to the expected 1. At the maximum τ=99, they are 3.2616e-2 and 2.9858e-2 respectively, quite close to the same, and have an Allan/ Thêo1 variance ratio of 1.1933, reasonably close to 1. The complete set of these values is shown below.
5. The computed $\Theta_1$ bias factor (as returned by the $\Theta_0\Theta_1BM$ function) for the 1001 phase data point test suite is 1.085666, consistent with its white FM noise type.

6. Simulated 1001-point phase data sets were generated for each of the five power law noise types, and the resulting $\Theta_1$ bias factors were compared with their original values as given in Reference [3], the range of their more accurate ones that have a slight dependence on $m$ as given in Reference [8] for $m=12$ thru 132, and the bias factor results of Stable32:

<table>
<thead>
<tr>
<th>Noise Type</th>
<th>W PM</th>
<th>F PM</th>
<th>W FM</th>
<th>F FM</th>
<th>RW FM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta_1$ Bias Factor</td>
<td>Original</td>
<td>0.4</td>
<td>0.6</td>
<td>1.0</td>
<td>1.71</td>
</tr>
<tr>
<td>Accurate</td>
<td>0.364-0.195</td>
<td>0.529-0.330</td>
<td>1.000</td>
<td>1.722-1.848</td>
<td>2.515-2.676</td>
</tr>
<tr>
<td>Stable32</td>
<td>0.191</td>
<td>0.356</td>
<td>0.733</td>
<td>1.600</td>
<td>1.854</td>
</tr>
<tr>
<td>Actual</td>
<td>0.255</td>
<td>0.428</td>
<td>1.180</td>
<td>1.515</td>
<td>2.339</td>
</tr>
</tbody>
</table>

There clearly is a wide variation in the bias factor results, but they all have the same general trend.

7. The edfs for setting $\Theta_0\Theta_1BM$ confidence limits are computed with the same C code as used in Stable32. They were verified as the same.

8. The Chi-squared values for setting single and double-sided error bars are computed using the $\texttt{qchisq()}$ R function. That function’s usage was confirmed against published tables and verified against the Stable32 values.

9. Selected single and double-sided $\Theta_0\Theta_1BM$ minimum, nominal, and maximum values were verified against Stable32.

10. Regardless of all of the above checks, this program is new and it should be used with care. Please report any problems to bill@wriley.com.
Appendix 4

R Top-Level Driver Function for a ThêoBR Run

With R, it is easy to write a top-level driver function for a ThêoBR run according to your preferences. This is especially convenient for processing phase data files of differing sizes, and to display and/or plot the analysis results. An example of such an R driver function is shown below:

The `trun()` “ThêoBR Run” function requires the `x` argument with the phase data array name (one assumes that it has already been read into R with, for example, `> x< scan("C:\Data\phase.dat")`). The `trun()` function has optional arguments for setting the data sampling interval and error bar confidence interval. It always sets `br=1` for bias removal and generates double-sided error bars. The ThêoBR results are plotted along with their error bar region, and a data frame with the tabulated tau, minimum, nominal, and maximum. The bias correction factor and estimated power law noise type exponent are printed, and the ThêoBR results are returned invisibly. To get those results, assign the function return to a variable name, e.g., `> r< trun(x)`. The `x` array is not changed and the `tau`, l, t, and u arrays are not returned or visible to the caller. The printed bias correction factor implies the dominant noise type. This driver function can be invoked on the R command line and saved as a `trun.R` file.

```r
trun <- function(x, tau0=1.0, ci=0.683) {
  # Allocate arrays
  N=length(x)
t<1:((N-1)/2)-1
u<1:((N-1)/2)-1
l<1:((N-1)/2)-1
tau<1:((N-1)/2)-1

  # Call TheoBR()
  cf=TheoBR(x,t,u,l,tau,tau0,ci,1)

  # Print the bias correction factor
  print(cf)

  # Crudely estimate noise type using original nominal cf
  # WPM=0.4, FPM=0.6, WFM=1.0, FFM=1.71, RWFM=2.24
  if(cf<sqrt(0.4*0.6)) a=2
  else if(cf<sqrt(0.6*1.0)) a=1
  else if(cf<sqrt(1.0*1.71)) a=0
  else if(cf<sqrt(1.71*2.24))a=-1
  else a=-2

  # Print power law noise exponent
  print(a)

  # Create data frame for results
  r< data.frame(tau,l,t,u)

  # Get plot y scale
  top=10^(ceiling(log10(max(t))))
  bot=10^(floor(log10(min(t))))

  # Plot results
  plot(tau,t,type="l",ylim=c(bot,top),main="Theo1BR Plot",xlab="Tau",ylab="Theo1BR",log="xy",col="red")
  polygon(c(tau, rev(tau)), c(u, rev(l)), col = "grey70", border = NA)
  points(tau,t,type="l",col="red")
  grid()
}
```
# Return data frame without printing it
invisible(r)
}

If desired, log scale ticks can then be added to the plot with another R function and example (not used in the following plots):

```r
logtics <- function(x, y) {
  # Get log scale limits
  left = floor(log10(min(x)))
  right = ceiling(log10(max(x)))
  bot = floor(log10(min(y)))
  top = ceiling(log10(max(y)))
  # Draw X-Axis Tics
  for(xx in left:right) {
    axis(side=1, at=(2:9)*10^xx, labels=FALSE)
  }
  # Draw Y-Axis Tics
  for(yy in bot:top) {
    axis(side=2, at=(2:9)*10^yy, labels=FALSE)
  }
}
```

This function was invoked for the 1001-point phase data test suite with the confidence interval set to 95% to emphasize the error bars. The resulting ThêoBR plot is shown below where the grey shading denotes the 2-sigma (95%) confidence interval region. More elaborate plot formats are, of course, possible. The ThêoBR bias factor was 1.085666 (W FM).
Create and display the first few rows of the data frame with the ThêoBR results (tau, minimum, nominal and maximum ThêoBR):

```r
data.frame(tau, l, t, u)
r
```

<table>
<thead>
<tr>
<th>tau</th>
<th>l</th>
<th>t</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.236361199</td>
<td>0.24861662</td>
<td>0.26222248</td>
</tr>
<tr>
<td>3.0</td>
<td>0.163755974</td>
<td>0.17243897</td>
<td>0.18210164</td>
</tr>
<tr>
<td>4.5</td>
<td>0.133973395</td>
<td>0.14162209</td>
<td>0.15020401</td>
</tr>
<tr>
<td>6.0</td>
<td>0.117754145</td>
<td>0.12502493</td>
<td>0.13325995</td>
</tr>
<tr>
<td>7.5</td>
<td>0.105102919</td>
<td>0.11208706</td>
<td>0.12007305</td>
</tr>
<tr>
<td>9.0</td>
<td>0.095475220</td>
<td>0.10225839</td>
<td>0.11008707</td>
</tr>
<tr>
<td>10.5</td>
<td>0.088026711</td>
<td>0.09467148</td>
<td>0.10240968</td>
</tr>
<tr>
<td>12.0</td>
<td>0.082062725</td>
<td>0.08860804</td>
<td>0.09629674</td>
</tr>
<tr>
<td>13.5</td>
<td>0.075792700</td>
<td>0.08214976</td>
<td>0.08967987</td>
</tr>
<tr>
<td>15.0</td>
<td>0.069693163</td>
<td>0.07581495</td>
<td>0.08312496</td>
</tr>
<tr>
<td>16.5</td>
<td>0.064784317</td>
<td>0.07072285</td>
<td>0.07786934</td>
</tr>
<tr>
<td>18.0</td>
<td>0.060899059</td>
<td>0.06670681</td>
<td>0.07374865</td>
</tr>
<tr>
<td>19.5</td>
<td>0.058012463</td>
<td>0.06375288</td>
<td>0.07076398</td>
</tr>
<tr>
<td>21.0</td>
<td>0.055112241</td>
<td>0.06075759</td>
<td>0.06770129</td>
</tr>
</tbody>
</table>
The same data set was analyzed in Stable32 which produced similar ThêoBR results.
Note that this is actually a “Many Tau” plot with N=500.
Appendix 5

Additional Examples

White PM  wpm.phd  wpm

Theo1BR Plot

> r<-trun(wpm)
> r

<table>
<thead>
<tr>
<th>tau</th>
<th>l</th>
<th>t</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.390243138</td>
<td>0.400441682</td>
<td>0.411483464</td>
</tr>
<tr>
<td>3.0</td>
<td>0.215129195</td>
<td>0.220876926</td>
<td>0.227111109</td>
</tr>
<tr>
<td>4.5</td>
<td>0.150534255</td>
<td>0.154860301</td>
<td>0.159581855</td>
</tr>
<tr>
<td>6.0</td>
<td>0.121114878</td>
<td>0.124874884</td>
<td>0.129008079</td>
</tr>
<tr>
<td>7.5</td>
<td>0.098608189</td>
<td>0.101898904</td>
<td>0.105542464</td>
</tr>
<tr>
<td>9.0</td>
<td>0.085601115</td>
<td>0.088609398</td>
<td>0.092009566</td>
</tr>
<tr>
<td>10.5</td>
<td>0.076077409</td>
<td>0.078954987</td>
<td>0.082185663</td>
</tr>
<tr>
<td>12.0</td>
<td>0.065410399</td>
<td>0.068021821</td>
<td>0.070973019</td>
</tr>
<tr>
<td>13.5</td>
<td>0.059140764</td>
<td>0.061621141</td>
<td>0.064442045</td>
</tr>
<tr>
<td>15.0</td>
<td>0.054863471</td>
<td>0.057270859</td>
<td>0.060025313</td>
</tr>
</tbody>
</table>
The ThêoBR and Stable32 Sigma functions produce the same nominal Thêo1 and bias factor for the WPM noise, and nearly the same 1-sigma error bars.
The Stable32 ThéoBR Run and Plot produce slightly different results because of a different bias factor.
> r <- trun(fpm)
> r

<table>
<thead>
<tr>
<th>tau</th>
<th>l</th>
<th>t</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5</td>
<td>0.505634967</td>
<td>0.518849140</td>
</tr>
<tr>
<td>2</td>
<td>3.0</td>
<td>0.315628560</td>
<td>0.324061390</td>
</tr>
<tr>
<td>3</td>
<td>4.5</td>
<td>0.233480032</td>
<td>0.240189769</td>
</tr>
<tr>
<td>4</td>
<td>6.0</td>
<td>0.189619689</td>
<td>0.195506425</td>
</tr>
<tr>
<td>5</td>
<td>7.5</td>
<td>0.162435703</td>
<td>0.167856446</td>
</tr>
<tr>
<td>6</td>
<td>9.0</td>
<td>0.142895464</td>
<td>0.147988126</td>
</tr>
<tr>
<td>7</td>
<td>10.5</td>
<td>0.127386652</td>
<td>0.132204967</td>
</tr>
<tr>
<td>8</td>
<td>12.0</td>
<td>0.113034665</td>
<td>0.117547423</td>
</tr>
<tr>
<td>9</td>
<td>13.5</td>
<td>0.104683239</td>
<td>0.109073678</td>
</tr>
<tr>
<td>10</td>
<td>15.0</td>
<td>0.094632242</td>
<td>0.098784669</td>
</tr>
</tbody>
</table>
The ThêoBR and Stable32 Sigma functions produce the same nominal Thêo1 and bias factor for the F PM noise, and nearly the same 1-sigma error bars. The noise is identified as W PM rather than F PM.
The Stable32 ThêoBR Run and Plot produce slightly different results because of a different bias factor.
> r<-trun(wfm)
> r

<table>
<thead>
<tr>
<th></th>
<th>tau</th>
<th>t</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5</td>
<td>0.86912455</td>
<td>0.89183809</td>
</tr>
<tr>
<td>2</td>
<td>3.0</td>
<td>0.61581492</td>
<td>0.63226801</td>
</tr>
<tr>
<td>3</td>
<td>4.5</td>
<td>0.49274597</td>
<td>0.50690647</td>
</tr>
<tr>
<td>4</td>
<td>6.0</td>
<td>0.4233857</td>
<td>0.43648110</td>
</tr>
<tr>
<td>5</td>
<td>7.5</td>
<td>0.37369928</td>
<td>0.38617023</td>
</tr>
<tr>
<td>6</td>
<td>9.0</td>
<td>0.34062912</td>
<td>0.35276883</td>
</tr>
<tr>
<td>7</td>
<td>10.5</td>
<td>0.31630424</td>
<td>0.32826824</td>
</tr>
<tr>
<td>8</td>
<td>12.0</td>
<td>0.29588356</td>
<td>0.30769631</td>
</tr>
<tr>
<td>9</td>
<td>13.5</td>
<td>0.27515789</td>
<td>0.28669808</td>
</tr>
<tr>
<td>10</td>
<td>15.0</td>
<td>0.25982648</td>
<td>0.27122757</td>
</tr>
</tbody>
</table>
The ThéoBR and Stable32 Sigma functions produce the same nominal Théo1 and bias factor for the WFM noise, nearly the same 1-sigma error bars.
The Stable32 ThêoBR Run and Plot produce slightly different results because of a different bias factor.

<table>
<thead>
<tr>
<th>AF</th>
<th>Thêo1 Tau</th>
<th>#</th>
<th>Alpha</th>
<th>Thêo1 Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>7.5000e+00</td>
<td>4955</td>
<td>0</td>
<td>3.5549e-01</td>
</tr>
<tr>
<td>12</td>
<td>9.0000e+00</td>
<td>5934</td>
<td>0</td>
<td>3.2474e-01</td>
</tr>
<tr>
<td>14</td>
<td>1.0500e+01</td>
<td>6909</td>
<td>0</td>
<td>3.0219e-01</td>
</tr>
<tr>
<td>16</td>
<td>1.2000e+01</td>
<td>7880</td>
<td>0</td>
<td>2.8325e-01</td>
</tr>
<tr>
<td>18</td>
<td>1.3500e+01</td>
<td>8847</td>
<td>0</td>
<td>2.6392e-01</td>
</tr>
<tr>
<td>20</td>
<td>1.5000e+01</td>
<td>9810</td>
<td>0</td>
<td>2.4968e-01</td>
</tr>
<tr>
<td>22</td>
<td>1.6500e+01</td>
<td>10769</td>
<td>0</td>
<td>2.4247e-01</td>
</tr>
<tr>
<td>24</td>
<td>1.8000e+01</td>
<td>11724</td>
<td>0</td>
<td>2.3647e-01</td>
</tr>
<tr>
<td>26</td>
<td>1.9500e+01</td>
<td>12675</td>
<td>0</td>
<td>2.2802e-01</td>
</tr>
<tr>
<td>28</td>
<td>2.1000e+01</td>
<td>13622</td>
<td>0</td>
<td>2.2241e-01</td>
</tr>
<tr>
<td>30</td>
<td>2.2500e+01</td>
<td>14585</td>
<td>0</td>
<td>2.1685e-01</td>
</tr>
<tr>
<td>32</td>
<td>2.4000e+01</td>
<td>15504</td>
<td>0</td>
<td>2.0797e-01</td>
</tr>
<tr>
<td>34</td>
<td>2.5500e+01</td>
<td>16439</td>
<td>0</td>
<td>2.0103e-01</td>
</tr>
<tr>
<td>36</td>
<td>2.7000e+01</td>
<td>17370</td>
<td>0</td>
<td>1.9442e-01</td>
</tr>
<tr>
<td>38</td>
<td>2.8500e+01</td>
<td>18297</td>
<td>0</td>
<td>1.8979e-01</td>
</tr>
</tbody>
</table>
Theo1BR Plot

> r<-trun(ffm)
> r
1.514804

> r

```r
tau l t u
1 1.5 1.5 1.1546793 1.1848555 1.2175267
2 3.0 0.9838932 1.0101804 1.0386924
3 4.5 0.9474265 0.9746336 1.0043700
4 6.0 0.9425662 0.9718281 1.0039944
5 7.5 0.9468058 0.9784023 1.0133866
6 9.0 0.9584494 0.9926077 1.0309666
7 10.5 0.9618735 0.9982557 1.0391023
8 12.0 0.9649376 1.0034615 1.0469977
9 13.5 0.9609765 1.0012801 1.0471169
10 15.0 0.9496628 0.9913337 1.0390156
```
The ThéoBR and Stable32 Sigma functions produce the same nominal Théo1 and bias factor for the FM noise, nearly the same 1-sigma error bars.
The Stable32 ThéoBR Run and Plot produce slightly different results because of a different bias factor.
Random Walk FM  rwfm.phd  rwfm

```
> r<-trun(rwfm)
[1] 2.339171
[1] -2

   tau  l  t  u
1  1.5 1.496699 1.535814 1.578162
2  3.0 1.829164 1.878034 1.931041
3  4.5 2.099929 2.160276 2.226141
4  6.0 2.343586 2.416342 2.496320
5  7.5 2.558800 2.644191 2.738738
6  9.0 2.747783 2.845712 2.954909
7 10.5 2.914554 3.024795 3.148563
8 12.0 3.060710 3.182905 3.320998
9 13.5 3.192064 3.325940 3.478196
10 15.0 3.309724 3.454953 3.621132
```
The ThêôBR and Stable32 Sigma functions produce the same nominal Thêô1 and bias factor for the RW FM noise, nearly the same 1-sigma error bars.
The Stable32 ThêoBR Run and Plot produce slightly different results because of a different bias factor.
Simulated Rubidium Frequency Standard Noise

10 Days (86,400 points) of 10-second W and F FM data:

\[ \sigma_y(\tau) = 1 \times 10^{-11} \tau^{-1/2} + 2 \times 10^{-13} \]

No Drift
> trun(rb,10)

[1] 1.250157

[1] 0
This basic 86,400 point All Tau Thêo1 calculation (no bias removal or error bars) in R using the Lewis fast algorithm takes about 5 seconds, while the same calculation using the naïve method in Stable32 takes 1317 seconds (22 minutes), about x260 times longer. An All Tau Thêo1 calculation takes longer than a ThêoH calculation because most of the points are ADEV values in the latter.

The otherwise same ThêoBR calculation and plot is nearly identical, with a Bias Factor of 1.039 ($\alpha=0$, W FM noise) and calculation time (1311s), and, for practical purposes, the same as the R results.
The 86,400-point ADEV calculation is practically instantaneous, the 8640-point x10 averaged ThéoH calculation takes a brief 9 seconds, and the unaveraged ThéoH calculation is calculation takes a bit over 11 minutes. The averaging does not affect the longer tau ThéoH results, which extend to about 600,000 seconds compared with about 200,000 seconds for ADEV.
Performing the un-averaged ThêoH run in Stable32 at octave tau increments takes only 20 seconds.
As a final example, this plot shows the results of a 1 million point, $\tau=1s$, $\sigma_y(1)=1.0$ ThêoBR run under R. The simulated white FM phase noise was generated in Stable32, and it took about 40 minutes to process, representing the practical size limit to this approach, one that would be totally impractical with the naïve algorithm. The plot shows that the instability falls somewhat faster that $\tau^{-1/2}$, and the analysis identifies it as closest to F PM noise.

The corresponding overlapping Allan deviation plot, shown below, is, of course, similar but covers a smaller tau range. It takes only about 7 seconds to compute in Stable32, so would be the analysis method of choice unless results were needed at large tau. If they were, the fast ThêoBR analysis is quite practical, especially compared with extending the data run.

The ADEV slopes show the noise to be W FM in the short term, becoming close to F FM and then F PM in the medium term, and tending toward F FM in the long term.

Thêo1 is known to be less prone to having oscillatory variations, but those are about the same as for ADEV in this run, so they are probably artifacts in the simulated data.

The same 1-sigma error bars in the octave ADEV plot are corresponding larger than those for the ThêoBR run, especially since those are for F FM rather than W FM noise.