Article

Accurate and Efficient Explicit Approximations of the Colebrook Flow Friction Equation Based on the Wright $\omega$-Function

Dejan Brkić 1,2,* and Pavel Praks 1,3,*

1 European Commission, Joint Research Centre (JRC), 21027 Ispra, Italy
2 IRC Alfatec, 18000 Niš, Serbia
3 IT4Innovations, VŠB—Technical University of Ostrava, 708 00 Ostrava, Czech Republic
* Correspondence: dejanbrkic0611@gmail.com (D.B.); pavel.praks@vsb.cz or pavel.praks@gmail.com (P.P.)

Received: 23 October 2018; Accepted: 24 December 2018; Published: 31 December 2018

Abstract: The Colebrook equation is a popular model for estimating friction loss coefficients in water and gas pipes. The model is implicit in the unknown flow friction factor, $f$. To date, the captured flow friction factor, $f$, can be extracted from the logarithmic form analytically only in the term of the Lambert $W$-function. The purpose of this study is to find an accurate and computationally efficient solution based on the shifted Lambert $W$-function also known as the Wright $\omega$-function. The Wright $\omega$-function is more suitable because it overcomes the problem with the overflow error by switching the fast growing term, $y = W(e^x)$, of the Lambert $W$-function to series expansions that further can be easily evaluated in computers without causing overflow run-time errors. Although the Colebrook equation transformed through the Lambert $W$-function is identical to the original expression in terms of accuracy, a further evaluation of the Lambert $W$-function can be only approximate. Very accurate explicit approximations of the Colebrook equation that contain only one or two logarithms are shown. The final result is an accurate explicit approximation of the Colebrook equation with a relative error of no more than 0.0096%. The presented approximations are in a form suitable for everyday engineering use, and are both accurate and computationally efficient.

Keywords: Colebrook equation; hydraulic resistance; Lambert $W$-function; Wright $\omega$-function; explicit approximations; computational burden; turbulent flow; friction factor

1. Introduction

The Colebrook equation; Equation (1), is an empirical relation which, in its native form, relates implicitly the unknown Darcy's flow friction factor, $f$, with the known Reynolds number, $R$, and the known relative roughness of inner pipe surface, $\epsilon^*$ [1,2]. Engineers use it at defined domains of the input parameters: $4000 < R < 10^8$ and for $0 < \epsilon^* < 0.05$. The Colebrook equation is transcendental (cannot be expressed in terms of elementary functions); the implicitly given function in respect to the unknown flow friction factor, $f$:

$$\frac{1}{\sqrt{f}} = -2 \cdot \log_{10} \left( \frac{2.51}{R} \cdot \frac{1}{\sqrt{f}} + \frac{\epsilon^*}{3.71} \right)$$

(1)

The Colebrook equation; Equation (1) also has an exact explicit analytical form in the terms of the Lambert $W$-function; Equation (2) [3,4], which is also transcendental, but which can be evaluated...
using numerous thoroughly tested procedures with varying accuracy and complexity, developed for various applications in physics and engineering [5].

\[
\frac{1}{\sqrt{f}} = \frac{2}{\ln(10)} \cdot \ln\left(\frac{R}{\sqrt{\ln(10)} \cdot \frac{\ln(10)}{2}}\right) + W(e^x) - x
\]

The parameter, \(x\), in Equation (2) depends on the input parameters; the Reynolds number, \(R\), and the relative roughness of the inner pipe surface, \(\varepsilon^*\). Its domain is \(7.51 < x < 618,187.84\). The Lambert \(W\)-based Colebrook equation; Equation (2), contains the fast growing term, \(W(e^x)\), which cannot be accurately stored in common computer registers due to the runtime overflow error for the particular combinations of the Reynolds number, \(R\), and the relative roughness of the inner pipe surface, \(\varepsilon^*\), that can easily occur in everyday engineering practice [6,7]. The problem can be solved using the Wright \(\omega\)-function, a cognate of the Lambert \(W\)-function, which uses a shifted, non fast-growing argument [4,8,9].

According to the very recently published study of Belkić [10], the Lambert \(W\)-function plays a very important role across interdisciplinary research. The reference gives an updated detailed list of applications, which include mathematics, physics, astrophysics, chemistry, biology, medicine, population genetics, ecology, sociology, education, energetics, and technology. Finally, the presented analytical solutions are numerically illustrated in the genome multiplicity corrections for survival of synchronous cell populations after irradiation [11].

This paper is oriented to the application of the Lambert \(W\)-function and the Wright \(\omega\)-function in hydraulics (fluid dynamics). We present a few approximate solutions of the transformed Lambert \(W\)-based Colebrook equation in a form more suitable for the computing codes used in various engineering software. All mentioned and developed approximations are summarized in the Appendix A. The best version of the presented explicit approximation gives the value of the flow friction factor, \(f\), for which the Colebrook equation is in balance with the relative error of no more than 0.0096\%. Such accuracy achieved without using a large number of computationally expensive logarithmic functions (or non-integer powers) is highly computationally efficient. As reported by Clamond [12], Winning and Coole [13], Biberg [4], Vatankhah [14], etc., functions, such as logarithms and non-integer powers, require special algorithms with the execution of many more floating-point operations compared with basic arithmetic operations (+, −, ×, /) that are executed directly in the central processor unit (CPU) of computers.

In our previous contributions, we accelerated and simplified an iterative solution of the Colebrook equation. In [15], the Padé approximation is used as a cheap alternative to the logarithm in the second and the successful iterations of the fixed-point method. In [16], low cost starting points for iterative methods found by genetic approximations of the Colebrook equation are introduced and tested, while Praks and Brkić [17] discusses the optimal multi-point iterative methods for the Colebrook equation. Finally, advanced iterative procedures for the Colebrook equation are studied in [18].

Rather than improving iterative solutions, in this paper, we provide simple, but powerful approximations of the exact solution of the Colebrook model, which is given by the Wright \(\omega\)-function. We introduce cheap yet still accurate approximations of the Wright \(\omega\)-function for the Colebrook model using a symbolic regression technique [16,19,20] and Padé approximation [15,21]. Apparently, this is the first highly accurate explicit approximation of the Colebrook equation that contains only two computationally expensive functions (two logarithms, or as an alternative, two functions with non-integer powers), or even less if a combination of Padé approximations [15,21,22], and symbolic regression is used for a further reduction of the computational burden (where one of the logarithms is approximated by simple rational functions with a moderate increase of the maximal relative error).
2. Proposed Explicit Approximations and Comparative Analysis

The Colebrook equation in the terms of the Lambert \( W \)-function was apparently first proposed in 1998 by Keady [3]. However, as confirmed by Sonnad and Goudar [6] and Brkić [7], the term, \( W(e^x) \), grows so fast that it cannot be evaluated easily, even in the registers of modern computers, due to the overflow runtime error for a particular number of combinations of the input parameters; the Reynolds number, \( R \), and the relative roughness of the inner pipe surface, \( \varepsilon \); where parameter \( x \) of Equation (2) depends directly on them. The procedure presented here replaces this fast growing term with the approximate form that can be used efficiently in everyday engineering practice; Equation (3):

\[
8.0878, \text{ transform the Lambert } W \text{-function,}
\]

\[
\omega \approx \ln(e^x) - \ln(\ln(e^x)) + \ln(\ln(e^x)) / \ln(e^x) \quad \text{where } x = \ln(e^x).
\]

Thus, the undesirably fast growing term, \( W(e^x) \), in Equation (2) is approximated accurately through \( y \approx x - \ln(x) + \ln(x) / x \). The transformation is based on unsigned Stirling numbers of the first kind as reported by Rollmann and Spindler [9]. Table 1 shows the values of \( W(e^x) \) compared with its approximate replacement in the domain of the applicability of the Colebrook equation. Without the proposed transformation and simplification, the runtime overflow error occurs during the evaluation of the friction factor, \( f \), in computers for particular pairs or the Reynolds number, \( R \), and the relative roughness of the inner pipe surface, \( \varepsilon \); where the parameter, \( x \), of Equation (2) depends directly on them (#VALUE! is an overflow error in Table 1). The values in Table 1 were calculated in MS Excel.

Table 1. Values of \( W(e^x) \) compared with its approximate replacement, \( y \approx x - \ln(x) + \ln(x) / x \).

<table>
<thead>
<tr>
<th>( W(e^x) )</th>
<th>( R = 4000 )</th>
<th>( R = 10^4 )</th>
<th>( R = 10^5 )</th>
<th>( R = 10^6 )</th>
<th>( R = 10^7 )</th>
<th>( R = 10^8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e^x = 10^{-6} )</td>
<td>5.763586714</td>
<td>6.552354737</td>
<td>8.994740889</td>
<td>10.78188015</td>
<td>13.94025768</td>
<td>26.71930109</td>
</tr>
<tr>
<td>( e^x = 10^{-5} )</td>
<td>5.767379666</td>
<td>6.562094186</td>
<td>8.694473428</td>
<td>11.80401384</td>
<td>24.50329461</td>
<td>125.7849498</td>
</tr>
<tr>
<td>( e^x = 10^{-3} )</td>
<td>5.803529409</td>
<td>6.695688386</td>
<td>9.67953496</td>
<td>22.29514802</td>
<td>124.50329461</td>
<td>125.7849498</td>
</tr>
<tr>
<td>( e^x = 10^{-2} )</td>
<td>6.186774452</td>
<td>7.63459358</td>
<td>20.0939172</td>
<td>122.325789</td>
<td>#VALUE!</td>
<td>#VALUE!</td>
</tr>
<tr>
<td>( e^x = 0.05 )</td>
<td>10.14320931</td>
<td>17.90904123</td>
<td>120.596067</td>
<td>#VALUE!</td>
<td>#VALUE!</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( y )</th>
<th>( R = 4000 )</th>
<th>( R = 10^4 )</th>
<th>( R = 10^5 )</th>
<th>( R = 10^6 )</th>
<th>( R = 10^7 )</th>
<th>( R = 10^8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e^x = 10^{-6} )</td>
<td>5.766606874</td>
<td>6.552971455</td>
<td>8.99238256</td>
<td>10.7784212</td>
<td>13.93654991</td>
<td>26.71669441</td>
</tr>
<tr>
<td>( e^x = 10^{-5} )</td>
<td>5.770385511</td>
<td>6.562602762</td>
<td>8.691991603</td>
<td>11.80037821</td>
<td>24.50049484</td>
<td>136.3596559</td>
</tr>
<tr>
<td>( e^x = 10^{-3} )</td>
<td>5.80193728</td>
<td>6.59024862</td>
<td>9.694862641</td>
<td>22.29214094</td>
<td>134.079366</td>
<td>1246.853296</td>
</tr>
<tr>
<td>( e^x = 10^{-2} )</td>
<td>6.188374207</td>
<td>7.633218988</td>
<td>20.0931698</td>
<td>124.552558</td>
<td>#VALUE!</td>
<td>#VALUE!</td>
</tr>
<tr>
<td>( e^x = 0.05 )</td>
<td>10.13993873</td>
<td>17.90960354</td>
<td>120.596067</td>
<td>124.552558</td>
<td>#VALUE!</td>
<td>#VALUE!</td>
</tr>
</tbody>
</table>

Note: #VALUE!—Overflow error.

The simplifications; \( W(e^x) - x \approx \ln(x) - (1/2 - 1); 2 / \ln(10) \approx 0.6866; 2.31 / \ln(10) \approx 2.18; \) and \( 2.18 - 3.71 \approx 8.0878 \), transform the Lambert \( W \)-based expression of the Colebrook equation in a very accurate explicit approximate form that can be used efficiently in everyday engineering practice; Equation (3):

\[
\begin{align*}
\frac{1}{\sqrt{f}} & \approx 0.8686 \left[ B - C + \frac{C}{\pi + A} \right] \\
A & \approx \frac{R e^x}{8.0878} \\
B & \approx \ln\left( \frac{R}{\ln R} \right) \approx \ln(R) - 0.779397488 \\
C & = \ln(B + A)
\end{align*}
\]

(3)
Instead of logarithmic functions in the proposed explicit approximation; Equation (3), a new form for \( B \) and \( C = \ln(B + A) \) can be introduced, where \( a \) can be any sufficiently large constant, where the larger value of \( a \) gives a more accurate approximation of the logarithmic function, Equation (4):

\[
\begin{align*}
B & \approx a \cdot R^{a^{-1}} - a - 0.779397488 \\
C & = \ln(B + A) \approx a \cdot (B + A)^{a^{-1}} - a 
\end{align*}
\]  

(4)

Very accurate results were obtained for \( a > 10^5 \). Choosing this value, power \( a^{-1} = \frac{1}{a} \) is a fraction with an integer numerator and denominator, where the appropriate form depends on the programming language, and the option with fewer floating point operations should be chosen [37].

Forms, such as \( R^{0.00001} \), require an evaluation of two transcendental functions, because compilers in most programming languages interpret it through \( e^{0.00001 \cdot \ln(R)} \) [12].

For more accurate results, \( W(e^x) - x \approx \frac{1.038 \cdot \ln(x)}{x + 0.332} - \ln(x) \) or \( W(e^x) - x \approx \frac{1.0119 \cdot \ln(x)}{x} - \ln(x) + \frac{\ln(x) - 2.3849}{x^2} \) can be used. These new approximations were found using the symbolic regression software, Eureqa [16,19,20], and they are 2.5 and 16.7 times more accurate, respectively, as compared with expression \( y \) from Table 1. The related approximations are given with Equations (5) and (6), respectively.

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \left[ B - C + \frac{1.038 \cdot C}{0.332 + B + A} \right] 
\]  

(5)

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \left[ B - C + \frac{1.0119 \cdot C}{B + A} + \frac{C - 2.3849}{(B + A)^2} \right] 
\]  

(6)

In Equations (5) and (6), parameters \( A, B, \) and \( C \) are the same as in Equation (3).

2.2. Accuracy

With the friction factor, \( f \), computed using the approximate equations, Equation (3), the Colebrook equation is in balance with a relative error of no more than 0.13%, while using Equation (5) of no more than 0.045%, and finally, using Equation (6) of no more than 0.0096%, respectively. The related distribution of errors is shown in Figure 1. The presented approximations require evaluation of only two computationally expensive functions (two logarithms; Equations (3), (5), and (6), or alternatively two non-integer powers; Equation (4)), and therefore they are not only accurate, but also efficient for calculation.

![Figure 1](image_url)  

Figure 1. Cont.
The demonstrated approximation of Equation (3) based on the Wright \(\omega\)-function with a relative error of up to 0.13% is about 10 times more accurate compared to the approximation from Brkić [30], while Equation (5) is more than 25 times more accurate than [30], and, finally, Equation (6) is more than 100 times more accurate than [30]. The approximations from Brkić [30,31] are based on the Lambert \(W\)-function.

The most accurate approximations available to date are by Vatankhah [14], Offor and Alabi [38], Buzzelli [39], Vatankhah and Kouchakzadeh [40], Romeo et al. [41], Zigrang and Sylvester [42], and Serghides [43]. All approximations mentioned [14,38–47] or developed in this paper are listed in the Appendix A, and their evaluated maximal relative error is shown in Table 2.

Elaboration on the accuracy of explicit approximations to the Colebrook equation can be found, among others, in Zigrang and Sylvester [48], Gregory and Fogarasi [49], Brkić [50,51], Winning and Coole [13,52], Brkić and Ćojbašić [11], and Pimenta et al. [53].

2.3. Complexity and Computational Burden

In a computer environment, a logarithmic function and non-integer powers require more floating-point operations to be executed in the central processor unit (CPU) compared to simple arithmetic operations, such as adding, subtracting, multiplication, and division [12–14,54,55]. With a relative error of up to 0.0096%, the herein proposed explicit approximation of the Colebrook equation; Equation (6), that contains only two computationally expensive functions, is not only...
accurate, but also sufficiently efficient. Winning and Coole [13] reported relative effort for computation as: Addition-1, Subtraction-1.18, Division-1.35, Multiplication-1.55, Squared-2.18, Square root-2.29, Cubed-2.38, Natural logarithm-2.69, Cubed root-2.71, Fractional exponential-3.32, and Logarithm to base 10–3.37. One exception is Biberg [4], who grouped division with more expensive functions.

Table 2. Number of computationally expensive functions in the available approximations of the Colebrook equations that introduce a relative error of no more than 1%.

<table>
<thead>
<tr>
<th>Approximation</th>
<th>Maximal Relative Error %</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Logarithms</td>
</tr>
<tr>
<td>Vatankhah [14]</td>
<td>0.0028%</td>
<td>1</td>
</tr>
<tr>
<td>Here developed; Equation (6)</td>
<td>0.0096%</td>
<td>2</td>
</tr>
<tr>
<td>Here developed; Equation (5)</td>
<td>0.045%</td>
<td>2</td>
</tr>
<tr>
<td>Ofior and Alabi [38]</td>
<td>0.0602%</td>
<td>2</td>
</tr>
<tr>
<td>Here developed; Equation (3)</td>
<td>0.13%</td>
<td>2</td>
</tr>
<tr>
<td>Here developed; Equation (4)</td>
<td>0.13%</td>
<td>0</td>
</tr>
<tr>
<td>Zigrang and Sylvester [42]</td>
<td>0.14%</td>
<td>3</td>
</tr>
<tr>
<td>Serghides [43]</td>
<td>0.14%</td>
<td>3</td>
</tr>
<tr>
<td>Romeo et al. [41]</td>
<td>0.14%</td>
<td>3</td>
</tr>
<tr>
<td>Vatankhah and Koukhakzadeh [40]</td>
<td>0.15%</td>
<td>2</td>
</tr>
<tr>
<td>Bar [41]</td>
<td>0.27%</td>
<td>2</td>
</tr>
<tr>
<td>Serghides-simple [43]</td>
<td>0.35%</td>
<td>2</td>
</tr>
<tr>
<td>Chen [45]</td>
<td>0.36%</td>
<td>2</td>
</tr>
<tr>
<td>Here developed; Equation (11)</td>
<td>Up to 0.4%</td>
<td>1</td>
</tr>
<tr>
<td>Fang et al. [46]</td>
<td>0.62%</td>
<td>1</td>
</tr>
<tr>
<td>Papaevangelou et al. [47]</td>
<td>0.82%</td>
<td>2</td>
</tr>
</tbody>
</table>

Notes: 1 All approximations are listed in the Appendix A of this paper, 2 in brackets: according to Clamond [12], non-integer powers require the evaluation of two computationally expensive functions–logarithm and exponential function, 3 in addition also contains one square root function.

For comparison, Table 2 provides the number of logarithmic functions and non-integer terms used in available approximations. Table 2 shows only highly accurate approximations with a relative error of no more than 1%, according to criteria set by Brkić [50]. All approximations from Table 2 are given in the Appendix A of this article.

In addition to that presented here, the approximation by Brkić [28,30,32] is also based on the Lambert W-function, but as it uses four logarithmic functions, it is much more computationally expensive, and is significantly less accurate with a relative error of about 2.2%.

In the next Section, the term, $B \approx \ln \left( \frac{R}{\frac{R}{T}} \right) \approx \ln(R) - 0.779397488$, from Equation (3) is approximated very accurately through rational polynomial expression, so complexity and computational expense additionally decrease, as the logarithm is accurately approximated by a simple rational function.

2.4. Simplifications

A simple rational approximation of the logarithm term, $B$, of the novel Colebrook approximation formulas; Equations (3), (5), and (6), is shown in this section. The logarithm represents the most computationally expensive operation of the Colebrook formula, see Equation (3). To reduce computation costs, the idea is to approximate the term, $B$, of Equation (3), which contains the logarithmic function, by a simple rational function. A combination of Padé approximation [15,21] and an artificial intelligence symbolic regression procedure [16,19,20] is used for this. Although the logarithm is a transcendental function, the found rational approximation remains simple and accurate with a maximal relative error limited to 0.2%. Although this rational approximation of the logarithm may seem unsightly to human eyes, it is very fast with computers, as it requires only a limited number of basic arithmetic operations to be executed in the central processor unit (CPU).
For the purpose of this simplification, the observed form, \( B \), from Equation (3) can be transformed as; Equation (7):

\[
B \approx \ln(R) - \ln(2.18) = \ln(315,012.6 \cdot r) - 0.77932 = \ln(r) + \ln(315,012.6) - 0.77932 = \ln(r) + 11.881
\]

(7)

In Equation (7), for the term, \( r = \frac{R}{315,012.6} \), the constant, 315,012.6, is carefully selected to minimize the error of the /2,3/ order Padé approximation of \( \ln(r) \) at the expansion point, \( r_0 = 1 \). The proposed Padé approximant of the /2,3/ order of \( \ln(r) \) at point 1 is; Equation (8):

\[
\ln(r) \approx s(r) = \frac{r \cdot (r(11\cdot r + 27) - 27) - 11}{r \cdot (r(3\cdot r + 27) + 27) + 3}
\]

(8)

The value, 315,012.6, is a weighted average of the Reynolds number, \( R \), for the turbulent zone valid for the Colebrook equation; \( R_{\text{min}} = 4000 \) and \( R_{\text{max}} = 10^8 \), using the value 0.0063 that was set by numerical experiments to minimize the absolute value of the maximum relative error of the Padé approximant of \( \ln(r) \) in the interval, \([R_{\text{min}}, R_{\text{max}}]\), as 0.0063. \( \frac{R_{\text{min}}+R_{\text{max}}}{2} = 315,012.6 \). The Padé approximant approximates a certain function very accurately in a relatively small domain of input parameters. It has been observed that the Padé approximant of \( \ln(r) \) at the expansion point, \( r_0 = 1 \), defined by a rational function, \( s(r) \), approximates \( \ln(r) \) with a maximal relative error of between \(-11.8\%\) and \(11.8\%\) for all values of the Reynolds number, \( R \), in the interval, \([R_{\text{min}}, R_{\text{max}}]\).

The Padé approximant, \( s(r) \), has a negligible error for \( r \sim 1 \), whereas top errors correspond to border points, \( R_{\text{min}} \) and \( R_{\text{max}} \). For example, for \( R_{\text{min}} = 4000 \), the Padé approximant of \( \ln(r) \) is \( s\left(\frac{4000}{315,012.6}\right) = s(0.012697905) = -3.38744594 \) where \( R = 315,012.6 \cdot r \). Therefore, the value of \( \ln(4000) \) is approximated by \( \ln(315,012.6) + (-3.38744549) = 9.27292448 \). The corresponding relative error for \( \ln(4000) \), is \(-11.8\%\).

Because of \( B = \ln\left(\frac{R}{100}\right) = \ln(R) - \ln\left(\frac{2.251}{\ln(10)}\right) \) and \( \ln(315,012.6) - \ln\left(\frac{2.251}{\ln(10)}\right) \sim 11.881 \), the value of \( B \) can be approximated as Equation (9):

\[
B \approx \ln(r) + 11.881 \approx s(r) + 11.881
\]

(9)

Further, a symbolic regression technique based on the computer software, Eureqa [19,20], is used for a more precise approximation of \( \ln(r) \). The aim is to construct a more accurate rational approximation of \( \ln(r) \) in comparison with Equation (9) using two known variables: The ratio, \( r = \frac{R}{315,012.6} \), and its Padé approximant, \( s(r) \). To reduce the burden for the central processor unit (CPU), the symbolic regression model should have a computationally cheap evaluation. For this reason, only rational functions are assumed for the symbolic regression model. To achieve that, 200 carefully selected quasi-random points of \( r \) using the LPTAU51 algorithm were used [56,57]. For these generated numbers, the Padé approximation \( s(r) \) was calculated using Equation (8). Also, \( \ln(r) \) was calculated to train the model in Eureqa for the purpose of finding a rational approximation of \( \ln(r) \) by using \( r \) and \( s(r) \) pairs. The developed models were successfully tested using 2048 quasi-random points. As a result, value \( B \) was approximated by simple rational functions, Equation (10), with a negligible maximal relative error of 0.0765%:

\[
B \approx 0.98236 \cdot s + \frac{s^7}{9200.67} + \frac{r}{150.2325} - \frac{r^2}{138187.1651} - \frac{1}{161.124} + 11.881
\]

(10)

Here, the symbol, \( s \), denotes the Padé approximant, \( s(r) \), given by Equation (8) and \( r = \frac{R}{315,012.6} \) is its argument.
When the Horner nested representation and the Variable Precision Arithmetic (VPA) at 4 decimal digit accuracy is assumed, the approximation of $B$ can be simplified by Equation (11):

$$B \approx s \cdot 0.0001086 \cdot 0.9824 - 0.006206 + 0.9824 - 0.000007237 \cdot r - 0.006656 + 11.88 \hspace{1cm} (11)$$

In this case, the maximal relative error remains negligible, 0.0793% compared with $B$ calculated using Equation (3).

The combined approach with the Padé approximant and the symbolic regression introduced in this section is based on human observation and introducing the ratio, $r = \frac{R}{315 \cdot 0.126}$, with the subsequent symbolic regression of $r$ and $s(r)$ pairs by Eureqa. The maximal relative error of $B$ introduced by Equation (11) was small (0.0793%), and in total if it is used instead of $B \approx \ln(\frac{R}{175})$ from Equation (3), the total maximal error of the explicit approximation of the Colebrook equation can go up to 0.4%. As can be seen from Table 2, with the only one-log call for $\ln(B + A)$ from Equation (3), this extremely accurate explicit approximation of the Colebrook equation approximation is the cheapest for computation presented to date.

The combined approach presented herein, with Padé approximation and symbolic regression, can be also used for faster yet still accurate probabilistic modelling of gas networks, which requires a large number of model evaluations [58–61].

### 3. Conclusions

In this paper, we presented novel, precise, but still numerically inexpensive approximations of the Colebrook flow friction equation, which has been widely used in hydraulics since 1937 [1–4]. Our novel approach exploited the fact that the exact solution of the Colebrook equation can be expressed by the Lambert $W$-function [3,4] and combined numerical properties of the Wright $\omega$-function [23–25], Padé approximation [15], and a symbolic regression technique [16,62].

The implicit Colebrook equation for flow friction is empirical, and hence has disputable accuracy. However, in many cases it is necessary to repeat calculations and to resolve the equation accurately to compare scientific results. An iterative solution [63] requires extensive computational effort, especially for flow evaluation of complex water or gas pipeline networks [64–67]. Although various available explicit approximations offer a good alternative, they are by the rule very accurate, but too complex, and vice versa [50]. Contrary to previous approximations of the Colebrook equation, the relation presented herein with a relative error limited to 0.0096% is amongst the most accurate available explicit approximations of the Colebrook equation. Moreover, the approach presented herein was also very computationally cheap, as it needs only one or two logarithms (or alternatively two non-integer powers). Measured against the criteria of accuracy and complexity, these approximations demonstrated desirable levels of performance. Consequently, the approximations can be recommended for implementation in software codes for engineering use.

The Colebrook equation is relevant only for turbulent flow, while for full-scale flow, different unified equations need to be used [68]. Consequently, the presented approximations are suitable for all cases where the Colebrook equation is used. Although the original experiment conducted by Colebrook and White used air [1,2], the Colebrook equation is today mostly used for modelling water flow [67]. However, in the case of natural gas flow modelling, the American Gas Association (AGA) and the American Bureau of Mines recommend replacing the coefficient of Equation (1) from 2.51 to 2.825 [69,70]. For future work, it would be interesting to integrate these novel flow friction models to stochastic gas network simulators [58,61].

Our approximations are valid for the whole domain of applicability of the Colebrook equation (on the other hand some of the approximations available from the scientific literature [71] are accurate only in certain narrow domain; e.g. for the highly turbulent flow or for the similar particular cases).
Author Contributions: The paper is a product of the joint efforts of the authors who worked together on models of natural gas distribution networks. D.B.’s scientific background is in control and applied computing in mechanical and petroleum engineering while P.P.’s in applied mathematics and programming. D.B. transformed the Colebrook equation through the Wright $\omega$-function and developed the first version of the approximations. P.P. worked on fine tuning of the model, and introduced Section 2.4. Simplifications. D.B. wrote the draft of this paper, and prepared figures and tables, while P.P. worked on its revision according to the reviewers’ reports.

Funding: Resources to cover the Article Processing Charge are provided by the European Commission. This communication is registered in the internal system for publication PUBSY of the Joint Research Centre of the European Commission under No. JRC113631.

Acknowledgments: We acknowledge support from the European Commission, Joint Research Centre (JRC), Directorate C: Energy, Transport and Climate, Unit C3: Energy Security, Distribution and Markets and we especially thank Marcelo Masera for his scientific supervision and approvals. This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia through the project iii44006 and by the Ministry of Education, Youth and Sports of the Czech Republic through the National Programme of Sustainability (NPS II) project “IT4Innovations excellence in science-LQ1602”. We also thank John Cawley, who as a native speaker kindly checked the correctness of English expressions throughout the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following symbols are used in this paper:

Constants:

$\alpha$ any > $10^5$

Variables:

$A$ variable that depends on $R$ and $\varepsilon^*$ (dimensionless)

$B$ variable that depends on $R$ (dimensionless)

$C$ variable that depends on variables $A$ and $B$ (dimensionless)

$f$ Darcy (Moody) flow friction factor (dimensionless)

$R$ Reynolds number (dimensionless)

$r$ variable that depends on $R$ (dimensionless)

$x$ variable in function on $R$ and $\varepsilon^*$ (dimensionless)

$\varepsilon^*$ Relative roughness of inner pipe surface (dimensionless)

$a$ variables defined in Appendix A of this paper

Functions:

$e$ exponential function

$log_{10}$ logarithm with base 10

$ln$ natural logarithm

$s$ Padé approximant

$W$ Lambert $W$-function

$\omega$ Wright $\omega$-function

Appendix A

The following explicit approximations of the Colebrook equation are referred to in this paper:

- Here, developed Equations (3), (5), and (6); Equations (A1)–(A3):

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \left[ B + C \left( \frac{1}{B + A} - 1 \right) \right] \quad (A1)
\]

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \left[ B + \frac{1.038 \cdot C}{0.332 + B + A} - C \right] \quad (A2)
\]

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \left[ B + \frac{1.0119 \cdot C}{B + A} - C + \frac{C - 2.3849}{(B + A)^2} \right] \quad (A3)
\]

where $A \approx \frac{R \cdot \varepsilon^*}{8.0878}$, $B \approx \ln \left( \frac{R}{2 \pi} \right) \approx \ln(R) - 0.779397488$, $C \approx \ln(B + A)$. 
Here, developed Equation (4); Equations (A4)–(A6):

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \left[ B + \left( a \cdot (B + A)^{a-1} - a \right) \left( \frac{1}{B + A} \right) - 1 \right] \tag{A4}
\]

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \left[ B + \frac{1.038 \left( a \cdot (B + A)^{a-1} - a \right)}{0.332 + B + A} - \left( a \cdot (B + A)^{a-1} - a \right) \right] \tag{A5}
\]

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \left[ B + \frac{1.019 \left( a \cdot (B + A)^{a-1} - a \right)}{1.0119 R} - \left( a \cdot (B + A)^{a-1} - a \right) + \frac{\left( a \cdot (B + A)^{a-1} - a \right) - 2.3849}{(B + A)^{a}} \right] \tag{A6}
\]

where \( A \approx \frac{R \epsilon}{1000}, \) and \( B \approx a \left( \frac{R}{2 \pi} \right)^{a-1} - a \approx a \cdot \left( R \right)^{a-1} - a - 0.779397488. \)

As parameter \( a \) is larger, the approximation is more accurate. The value, \( a > 10^5 \), gives the sufficiently accurate approximation for gas hydraulic modelling, as the corresponding maximal relative error is less than 0.007% for the analysed Colebrook model.

Here, developed Equation (11); Equation (A7):

Parameter \( B \) from the Equations (A1)–(A3) and Equations (A4)–(A6) should be calculated using Equation (A7):

\[
B \approx s \cdot \left( 0.0001086 \cdot s^6 + 0.9824 \right) - \frac{0.006326}{r} - \frac{0.000007237 \cdot r}{r - 0.006656} + 11.881 \tag{A7}
\]

- Buzzelli [39]; (A8):

\[
\frac{1}{\sqrt{f}} \approx a_1 - \left( a_{12} + 2 \log_{10} \left( \frac{2R}{4} \right) \right) \tag{A8}
\]

\[
a_1 \approx \frac{0.0074 \ln(R)}{1 + 1.32 \sqrt{r}} \tag{A8}
\]

\[
a_2 \approx \frac{2.51 \cdot 2}{\sqrt{R}} \tag{A8}
\]

- Zigrang and Sylvester [42]; (A9):

\[
\frac{1}{\sqrt{f}} \approx -2 \cdot \log_{10} \left( \frac{\epsilon}{17} - \frac{5.02}{R} \right) \tag{A9}
\]

\[
a_3 \approx \log_{10} \left( \frac{\epsilon}{17} - \frac{5.02}{R} \right) \tag{A9}
\]

\[
a_4 \approx \log_{10} \left( \frac{\epsilon}{17} - \frac{13}{R} \right) \tag{A9}
\]

- Serghides [43]; (A10):

\[
\frac{1}{\sqrt{f}} \approx a_5 - \frac{(a_4 - a_3)^2}{a_{52} + a_3 + a_5} \tag{A10}
\]

\[
a_5 \approx -2 \cdot \log_{10} \left( \frac{\epsilon}{17} - \frac{12}{R} \right) \tag{A10}
\]

\[
a_6 \approx -2 \cdot \log_{10} \left( \frac{\epsilon}{17} - \frac{13}{R} \right) \tag{A10}
\]

\[
a_7 \approx -2 \cdot \log_{10} \left( \frac{\epsilon}{17} - \frac{2.51}{R} \right) \tag{A10}
\]

- Romeo et al. [41]; (A11):

\[
\frac{1}{\sqrt{f}} \approx -2 \cdot \log_{10} \left( \frac{\epsilon}{3.005} - \frac{5.0272}{R} \cdot a_8 \right) \tag{A11}
\]

\[
a_8 \approx \log_{10} \left( \frac{\epsilon}{3.005} - \frac{5.0272}{R} \cdot a_9 \right) \tag{A11}
\]

\[
a_9 \approx \log_{10} \left( \frac{\epsilon}{7.7918} \cdot \left( \frac{0.9924}{1.0119} \cdot 5.3326 \cdot 0.0815 \cdot 0.9345 \right) \right) \tag{A11}
\]
- Vatankhah and Kouchakzadeh [40]; (A12):

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \cdot \ln \left( \frac{0.4587 \cdot R}{(a_{10} - 0.33175)} \right) \\
\alpha_{10} \approx 0.124 \cdot R \cdot e^\alpha + \ln(0.1587 \cdot R) \\
\alpha_{11} \approx \frac{a_{10}}{a_{10} + 0.05683}
\]

(A12)

- Barr [44]; (A13):

\[
\frac{1}{\sqrt{f}} \approx -2 \cdot \log_{10} \left( \frac{c^2 + 4.518 \log_{10}(\frac{c}{R})}{R} \right) \\
\alpha_{12} \approx R \cdot \left( 1 + \frac{0.83}{2\sqrt{R}} \cdot (e^\alpha)^{0.7} \right)
\]

(A13)

- Serghides-simple [43]; (A14):

\[
\frac{1}{\sqrt{f}} \approx 4.781 - \left( \frac{(\alpha_{13} - 4.781)^2}{e_{14} - 1.613 - 0.2479} \right)^2 \\
\alpha_{13} \approx -2 \cdot \log_{10} \left( \frac{c^2}{R} - \frac{12}{R} \right) \\
\alpha_{14} \approx -2 \cdot \log_{10} \left( \frac{c^2}{R} - \frac{22.1}{R} \cdot \alpha_{13} \right)
\]

(A14)

- Chen [45]; (A15):

\[
\frac{1}{\sqrt{f}} \approx -2 \cdot \log_{10} \left( \frac{c^2}{R} - \frac{5.052}{R} \cdot \alpha_{15} \right) \\
\alpha_{15} \approx \log_{10} \left( \frac{c^2}{R} - \frac{5.806}{R} \cdot \alpha_{15} \right)
\]

(A15)

- Fang et al. [46]; (A16):

\[
\frac{1}{\sqrt{f}} \approx (1.613 - (\ln(0.234 \cdot (\alpha_{16})^{1.1007} - \alpha_{16}))^{-2})^{-2} \\
\alpha_{16} \approx \frac{66.525}{10^{0.0001}} + \frac{56.291}{10^{0.0001}}
\]

(A16)

- Papaevangelou et al. [47]; (A17):

\[
\frac{1}{\sqrt{f}} \approx \left( \frac{0.2479 - 0.0000947 \cdot (7 - \log_{10}(R))^4}{\log_{10} \left( \frac{c^2}{R} + \frac{0.766}{10^{0.0001}} \right)} \right)^{-2}
\]

(A17)

- Vatankhah [14]; (A18):

\[
\frac{1}{\sqrt{f}} \approx 0.8686 \cdot \ln \left( \frac{0.3984 \cdot R}{(0.8668 \cdot a_{17})^{0.9171 \cdot R}} \right) \\
\alpha_{17} \approx 0.12363 \cdot R \cdot e^\alpha + \ln(0.3984 \cdot R) \\
\alpha_{18} \approx 1 + \frac{1}{\ln \left( \frac{0.5000 + a_{17}}{1 + a_{17}} \right)}
\]

(A18)

- Offor and Alabi [38]; (A19):

\[
\frac{1}{\sqrt{f}} \approx -2 \cdot \log_{10} \left( \frac{c^2}{R} - \frac{1.975}{R} \cdot \alpha_{19} \right) \\
\alpha_{19} \approx \ln \left( \frac{c^2}{R} \right) + \frac{0.992}{10^{0.0001}} + \frac{7.627}{10^{0.0001}} + \frac{395.9}{R^{0.0001}}
\]

(A19)

References
10. Belkić, D. All the trinomial roots, their powers and logarithms from the Lambert series, Bell polynomials and Fox–Wright function: Illustration for genome multiplicity in survival of irradiated cells. J. Math. Chem. 2018. [CrossRef]
17. Praks, P.; Brkić, D. Choosing the optimal multi-point iterative method for the Colebrook flow friction equation. Processes 2018, 6, 130. [CrossRef]


36. Mikata, Y.; Walczak, W.S. Exact analytical solutions of the Colebrook-White equation. *J. Hydraul. Eng.* 2016, 142, 04015050. [CrossRef]


42. Zigrang, D.J.; Sylvester, N.D. Explicit approximations to the solution of Colebrook’s friction factor equation. *AIChE J.* 1982, 28, 514–515. [CrossRef]


52. Winning, H.K.; Coole, T. Explicit friction factor accuracy and computational efficiency for turbulent flow in pipes. *Flow Turbul. Combust.* 2013, 90, 1–27. [CrossRef]


55. Taherdangkoo, R.; Taherdangkoo, M. Modified stem cells algorithm-based neural network applied to bottom hole circulating pressure in underbalanced drilling. *Int. J. Pet. Eng.* 2015, 1, 178. [CrossRef]

57. Bratley, P.; Fox, B.L. Algorithm 659: Implementing Sobol’s quasirandom sequence generator. *ACM Trans. Math. Softw. (TOMS)* 1988, 14, 88–100. [CrossRef]


60. Pambour, K.A.; Cakir Erdener, B.; Bolado-Lavin, R.; Dijkema, G.P. Development of a simulation framework for analyzing security of supply in integrated gas and electric power systems. *Appl. Sci.* 2017, 7, 47. [CrossRef]

