Making the Most of Parameter Estimation: Terpolymerization Troubleshooting Tips

Authors:

Alison J. Scott, Vida A. Gabriel, Marc A. Dubé, Alexander Penlidis

Abstract:

Multi-component polymers can provide many advantages over their homopolymer counterparts. Terpolymers are formed from the combination of three unique monomers, thus creating a new material that will exhibit desirable properties based on all three of the original comonomers. To ensure that all three comonomers are incorporated (and to understand and/or predict the degree of incorporation of each comonomer), accurate reactivity ratios are vital. In this study, five terpolymerization studies from the literature are revisited and the 'ternary' reactivity ratios are re-estimated. Some recent studies have shown that binary reactivity ratios (that is, from the related copolymer systems) do not always apply to ternary systems. In other reports, binary reactivity ratios are in good agreement with terpolymer data. This investigation allows for the comparison between previously determined binary reactivity ratios and newly estimated 'ternary' reactivity ratios for several systems. In some of the case studies presented herein, reactivity ratio estimation directly from terpolymerization data is limited by composition restrictions or ill-conditioned systems. In other cases, we observe similar or improved prediction performance (for ternary systems) when 'ternary' reactivity ratios are estimated directly from terpolymerization data (compared to the traditionally used binary reactivity ratios). In order to demonstrate the advantages and challenges associated with 'ternary' reactivity ratio estimation, five case studies are presented (with examples and counter-examples) and troubleshooting suggestions are provided to inform future work.

Making the Most of Parameter Estimation: Terpolymerization Troubleshooting Tips

Alison J. Scott [1] **, Vida A. Gabriel** [2] **, Marc A. Dubé** [2] **and Alexander Penlidis** [1,*]

1 Department of Chemical Engineering, Institute for Polymer Research, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada
2 Department of Chemical and Biological Engineering, Centre for Catalysis Research and Innovation, University of Ottawa, Ottawa, ON K1N 6N5, Canada
* Correspondence: penlidis@uwaterloo.ca

Abstract: Multi-component polymers can provide many advantages over their homopolymer counterparts. Terpolymers are formed from the combination of three unique monomers, thus creating a new material that will exhibit desirable properties based on all three of the original comonomers. To ensure that all three comonomers are incorporated (and to understand and/or predict the degree of incorporation of each comonomer), accurate reactivity ratios are vital. In this study, five terpolymerization studies from the literature are revisited and the 'ternary' reactivity ratios are re-estimated. Some recent studies have shown that binary reactivity ratios (that is, from the related copolymer systems) do not always apply to ternary systems. In other reports, binary reactivity ratios are in good agreement with terpolymer data. This investigation allows for the comparison between previously determined binary reactivity ratios and newly estimated 'ternary' reactivity ratios for several systems. In some of the case studies presented herein, reactivity ratio estimation directly from terpolymerization data is limited by composition restrictions or ill-conditioned systems. In other cases, we observe similar or improved prediction performance (for ternary systems) when 'ternary' reactivity ratios are estimated directly from terpolymerization data (compared to the traditionally used binary reactivity ratios). In order to demonstrate the advantages and challenges associated with 'ternary' reactivity ratio estimation, five case studies are presented (with examples and counter-examples) and troubleshooting suggestions are provided to inform future work.

Keywords: copolymerization; design of experiments; reactivity ratio estimation; terpolymerization

1. Introduction

Terpolymer systems are becoming increasingly popular in a variety of applications, as they allow for the combination of various desirable properties into a single material [1,2]. The diversity of monomers allows for the production of countless new multi-component materials, all with particular combinations of desirable properties for custom applications. In the past decade alone, terpolymers have been investigated for use as electronic materials [3–5], biomedical materials [6–8], pressure sensitive adhesives [9–12], coatings [13–15], and for many more applications. A good understanding of terpolymerization kinetics (for each unique terpolymer system) ensures that appropriate formulations are selected to achieve desirable properties in the product terpolymer. Namely, 'ternary' reactivity ratios (that is, reactivity ratios estimated directly from terpolymerization data using an appropriate model, as opposed to reactivity ratios estimated from the corresponding binary copolymerizations, alternatively referred to as 'ternary copolymerizations') provide information about the degree of incorporation of each comonomer into the product terpolymer (based on the initial pre-polymer

formulation). Therefore, we can select appropriate feed compositions to achieve a desirable terpolymer composition and microstructure [16–19].

Recent studies have shown that binary (copolymerization) reactivity ratios (estimated from copolymerization data) do not always apply to terpolymerizations [18,20]. Thus, in general, 'ternary' reactivity ratio estimation (directly from terpolymerization data) is recommended. 'Ternary' reactivity ratio estimation (estimating all six related reactivity ratios simultaneously from terpolymerization data using the recast Alfrey–Goldfinger model [18]) ensures that any interactions and contributions from the third comonomer are considered, which improves the kinetic and statistical accuracy of parameter estimates. When terpolymerization experiments are selected with parameter estimation in mind (that is, using an error-in-variables model (EVM)-based design of experiments for ternary systems [21]), it can be straightforward to estimate 'ternary' reactivity ratios using the information-rich dataset. However, in using historical data (which may not be designed for parameter estimation), experimental terpolymerization data may not be sufficient for analysis. Common limitations in terpolymerization studies include composition restrictions or a lack of experimental information content (minimal replication, formulations selected targeting final properties rather than information collection, etc.). Since working with historical (previously collected) data has some challenges, the case studies presented herein address those challenges and aim to improve 'ternary' reactivity ratio prediction to the extent which is possible. For example, additional targeted experiments (i.e., the sequential selection of experiments based on an existing dataset), experimental replication, and full-conversion (cumulative) analysis can supplement a pre-existing terpolymerization dataset. While additional targeted experiments and replication may require revisiting the lab (or simulating additional data), full-conversion data are often already available from earlier runs. Cumulative analysis can provide greater information content from fewer experimental runs [22].

Despite the fact that binary reactivity ratios are only a numerical approximation for ternary systems, several cases have reported that model predictions using binary reactivity ratios seem to be in good agreement with terpolymerization data (most recently [12]). In these cases, although binary reactivity ratios provide good prediction performance, there are still benefits associated with using 'ternary' reactivity ratios directly [20]. First, this ensures that no unfounded assumptions are made about the nature of the terpolymerization, where the addition of the third comonomer (and any interactions it might have with the other two comonomers) is carefully evaluated. Second, it reduces the experimental load required to characterize the system (e.g., two designed formulations of three copolymers require six experiments for binary reactivity ratios, compared to three designed formulations per single terpolymer requiring three experiments for 'ternary' reactivity ratios). This is especially important if a new (unknown) combination of comonomers is being investigated. Third, and perhaps most importantly, the kinetic and statistical accuracy of using 'ternary' reactivity ratio estimation in terpolymerization studies ensures accurate model predictions for the system being studied.

In the present work, several case studies are revisited to explore the advantages (and challenges) of directly estimating 'ternary' reactivity ratios from terpolymerization data. In the first two case studies (Section 3.1), common challenges in parameter estimation are addressed, such as process constraints (experimental/composition limitations) and numerical estimation constraints (ill-conditioned systems). Additional case studies (Section 3.2) are then presented to provide a direct comparison between binary and 'ternary' reactivity ratio estimates. When the above-mentioned challenges are not a factor, the advantages of estimating 'ternary' reactivity ratios are evident.
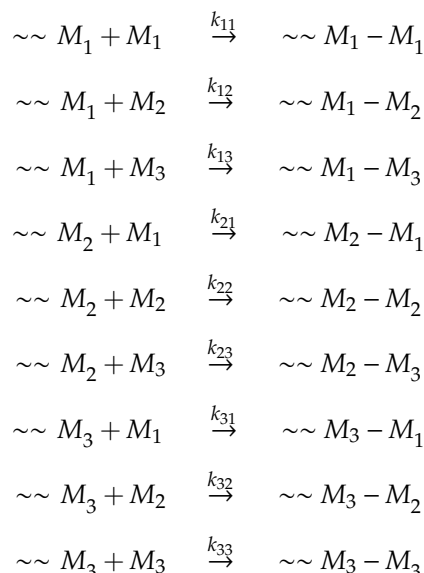
2. Background

2.1. Free Radical Terpolymerization

Terpolymerization systems (and multicomponent polymerizations in general) are part of an interesting and growing area of research, since there are countless combinations of monomers to be

discovered. However, because of the wide range of possibilities, these more complex terpolymerization systems have not been studied as thoroughly as copolymerization systems.

The kinetics of terpolymerization systems were first described by Alfrey and Goldfinger [23]. Given that there are three different possibilities for the terminal monomer (on the growing radical), and three options for the added monomer, nine different propagation steps are possible according to the terminal model:

$$\sim\sim M_1 + M_1 \xrightarrow{k_{11}} \sim\sim M_1 - M_1$$

$$\sim\sim M_1 + M_2 \xrightarrow{k_{12}} \sim\sim M_1 - M_2$$

$$\sim\sim M_1 + M_3 \xrightarrow{k_{13}} \sim\sim M_1 - M_3$$

$$\sim\sim M_2 + M_1 \xrightarrow{k_{21}} \sim\sim M_2 - M_1$$

$$\sim\sim M_2 + M_2 \xrightarrow{k_{22}} \sim\sim M_2 - M_2$$

$$\sim\sim M_2 + M_3 \xrightarrow{k_{23}} \sim\sim M_2 - M_3$$

$$\sim\sim M_3 + M_1 \xrightarrow{k_{31}} \sim\sim M_3 - M_1$$

$$\sim\sim M_3 + M_2 \xrightarrow{k_{32}} \sim\sim M_3 - M_2$$

$$\sim\sim M_3 + M_3 \xrightarrow{k_{33}} \sim\sim M_3 - M_3$$

The standard Alfrey–Goldfinger (AG) equations use ratios of instantaneous mole fractions in the terpolymer (i.e., F_i/F_j) as responses. However, measurements taken from experimental work are typically single mole fractions, not ratios, which means that evaluating (F_i/F_j) ratios results in lost information and a distorted error structure [18]. Thus, the AG model was re-derived (based on [18]) from basic terpolymerization kinetics for this study, so that each terpolymer mole fraction is presented as a single response (see Equations (1)–(3); details for this derivation are available upon request). This formulation is an improvement over the original AG model, as it promotes symmetrical estimation and error structures are not distorted. It also agrees with the recent work published by Kazemi et al. [18], in which the 'recast' version of the AG model was developed using the alternative ratio-based equations as a starting point.

$$F_1 - \frac{f_1\left(\frac{f_1}{r_{21}r_{31}} + \frac{f_2}{r_{21}r_{32}} + \frac{f_3}{r_{31}r_{23}}\right)\left(f_1 + \frac{f_2}{r_{12}} + \frac{f_3}{r_{13}}\right)}{f_1\left(\frac{f_1}{r_{21}r_{31}} + \frac{f_2}{r_{21}r_{32}} + \frac{f_3}{r_{31}r_{23}}\right)\left(f_1 + \frac{f_2}{r_{12}} + \frac{f_3}{r_{13}}\right) + f_2\left(\frac{f_1}{r_{12}r_{31}} + \frac{f_2}{r_{12}r_{32}} + \frac{f_3}{r_{13}r_{32}}\right)\left(f_2 + \frac{f_1}{r_{21}} + \frac{f_3}{r_{23}}\right) + f_3\left(\frac{f_1}{r_{13}r_{21}} + \frac{f_2}{r_{23}r_{12}} + \frac{f_3}{r_{13}r_{23}}\right)\left(f_3 + \frac{f_1}{r_{31}} + \frac{f_2}{r_{32}}\right)} = 0 \quad (1)$$

$$F_2 - \frac{f_2\left(\frac{f_1}{r_{12}r_{31}} + \frac{f_2}{r_{12}r_{32}} + \frac{f_3}{r_{13}r_{32}}\right)\left(f_2 + \frac{f_1}{r_{21}} + \frac{f_3}{r_{23}}\right)}{f_1\left(\frac{f_1}{r_{21}r_{31}} + \frac{f_2}{r_{21}r_{32}} + \frac{f_3}{r_{31}r_{23}}\right)\left(f_1 + \frac{f_2}{r_{12}} + \frac{f_3}{r_{13}}\right) + f_2\left(\frac{f_1}{r_{12}r_{31}} + \frac{f_2}{r_{12}r_{32}} + \frac{f_3}{r_{13}r_{32}}\right)\left(f_2 + \frac{f_1}{r_{21}} + \frac{f_3}{r_{23}}\right) + f_3\left(\frac{f_1}{r_{13}r_{21}} + \frac{f_2}{r_{23}r_{12}} + \frac{f_3}{r_{13}r_{23}}\right)\left(f_3 + \frac{f_1}{r_{31}} + \frac{f_2}{r_{32}}\right)} = 0 \quad (2)$$

$$F_3 - \frac{f_3\left(\frac{f_1}{r_{13}r_{21}} + \frac{f_2}{r_{23}r_{12}} + \frac{f_3}{r_{13}r_{23}}\right)\left(f_3 + \frac{f_1}{r_{31}} + \frac{f_2}{r_{32}}\right)}{f_1\left(\frac{f_1}{r_{21}r_{31}} + \frac{f_2}{r_{21}r_{32}} + \frac{f_3}{r_{31}r_{23}}\right)\left(f_1 + \frac{f_2}{r_{12}} + \frac{f_3}{r_{13}}\right) + f_2\left(\frac{f_1}{r_{12}r_{31}} + \frac{f_2}{r_{12}r_{32}} + \frac{f_3}{r_{13}r_{32}}\right)\left(f_2 + \frac{f_1}{r_{21}} + \frac{f_3}{r_{23}}\right) + f_3\left(\frac{f_1}{r_{13}r_{21}} + \frac{f_2}{r_{23}r_{12}} + \frac{f_3}{r_{13}r_{23}}\right)\left(f_3 + \frac{f_1}{r_{31}} + \frac{f_2}{r_{32}}\right)} = 0 \quad (3)$$

The newly derived equations are only valid for the instantaneous case. Using low conversion data makes it possible to assume that terpolymer composition drift is negligible (that is, at low conversions, the measurable cumulative copolymer composition is approximately equal to its instantaneous value). However, this restrictive assumption introduces additional sources of error, including significant experimental difficulties.

As an alternative, a cumulative ternary composition model can be employed to estimate 'ternary' reactivity ratios using the full conversion trajectory. The cumulative model shown in Equations (4)–(6) relates the cumulative terpolymer composition for each comonomer ($\overline{F}_i$) to the mole fraction of monomer

in the initial feed ($f_{i,0}$), the mole fraction of unreacted monomer in the polymerizing mixture (f_i), and the overall molar conversion (X_n).

$$\overline{F}_1 = \frac{f_{1,0} - f_1(1 - X_n)}{X_n} \tag{4}$$

$$\overline{F}_2 = \frac{f_{2,0} - f_2(1 - X_n)}{X_n} \tag{5}$$

$$\overline{F}_3 = \frac{f_{3,0} - f_3(1 - X_n)}{X_n} \tag{6}$$

If we cannot assume constant composition (that is, if composition drift is no longer negligible), f_i must be evaluated over conversion X_n, as shown in Equations (7)–(9).

$$\frac{df_1}{dX_n} = \frac{f_1 - F_1}{1 - X_n} \tag{7}$$

$$\frac{df_2}{dX_n} = \frac{f_2 - F_2}{1 - X_n} \tag{8}$$

$$\frac{df_3}{dX_n} = \frac{f_3 - F_3}{1 - X_n} \tag{9}$$

The models used to describe terpolymerization processes are simply extensions of the copolymer case, though slightly more complex. It is important to realize that ternary systems are unique and should not simply be viewed as a consolidation of the three analogous copolymer systems. When a third monomer is added to a polymerization formulation it can alter the polymerization characteristics and affect the degree of incorporation of all three comonomers [20]. Thus, whenever possible, terpolymerization kinetics should be evaluated and investigated in their own right.

The reader should note here from the outset that the term "'ternary' reactivity ratio" (used throughout this paper) refers to the collection of the six reactivity ratios that exist in Equations (1)–(3) (r_{12}, r_{13}, r_{21}, r_{23}, r_{31}, r_{32}). Although each individual parameter represents a relationship between two of the three comonomers, all six are estimated simultaneously from terpolymerization data, and therefore have an indirect influence on each other. For example, even though $r_{12} = k_{11}/k_{12}$ (and therefore only represents the propagation relationships for comonomers 1 and 2), it may still be influenced by the presence of monomer 3 in the polymerizing mixture. One may argue that since only two monomers are involved in the definition of each reactivity ratio, they are still technically 'binary' in nature. However, the term "'ternary' reactivity ratio" has been employed herein to emphasize that these reactivity ratios are estimated directly from terpolymerization data using the recast Alfrey–Goldfinger model. This is discussed further in what follows. It is our hope that creating this 'binary'/'ternary' distinction (related to both the source data and the series of equations employed for estimation) has simplified the reading and understanding of this work.

2.2. Reactivity Ratio Estimation

For the case studies that follow, the error-in-variables-model (EVM) was employed for the estimation of reactivity ratios. Through the EVM and direct numerical integration, we are able to estimate reactivity ratios using cumulative composition data (where the experimental data are available) [22]. This analysis provides additional advantages, including eliminating unnecessary assumptions and avoiding the experimental challenges associated with collecting low-conversion data.

The EVM methodology is an established tool for both selecting optimal feed compositions for reactivity ratio estimation experiments and for estimating reactivity ratios using non-linear parameter estimation. Therefore, only a brief overview is presented in what follows.

2.2.1. Design of Experiments for 'Ternary' Reactivity Ratio Estimation

When a series of experiments is designed in an optimal way, it becomes possible to minimize the number of experiments while increasing the information content from those experiments. Optimally designed experiments typically have much smaller joint confidence regions (JCRs), which are indicative of higher precision reactivity ratio estimates [24].

The EVM considers error in all variables, thus, using an experiment design technique within the EVM context helps to account for the error in both the independent variables (feed compositions) and the dependent variables (terpolymer compositions). The EVM can also take into account any experimental limitations, which ensures that the mathematical results of the experimental design are physically viable [21]. Of particular note is a practical heuristic for designing experiments for 'ternary' reactivity ratio estimation, which suggests that (statistically speaking) the optimal feed compositions always fall into the same range. When the terminal model is valid (as per Section 2.1), a multiplicative error structure is assumed and there are no other process constraints. The optimal feed compositions are as shown (in the shaded areas) in Figure 1. Three terpolymerization formulations, each rich in one comonomer, are sufficient to estimate 'ternary' reactivity ratios [21]. In a recent application, this was further confirmed while investigating the terpolymerization kinetics of 2-acrylamido-2-methylpropane sulfonic acid, acrylamide, and acrylic acid [20].
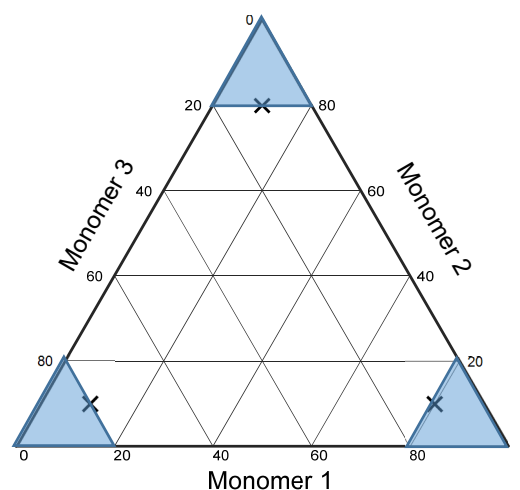

Figure 1. Optimally designed experiments for 'ternary' reactivity ratio estimation (as per [21]).

2.2.2. Reactivity Ratio Estimation Using EVM

The implementation of the EVM is based on Equations (10) and (11). Equation (10) relates the vector of known measurements ($\underline{x}_i$) to the vector of their unknown true values ($\underline{\xi}_i$) and an error term, $k\underline{\varepsilon}_i$ (where k is a constant that reflects the magnitude of measurement uncertainty, usually estimated from process information). Equation (11) is the model for the system, which shows the relationship between the true values of the variables ($\underline{\xi}_i$) and the true (but unknown) parameter values to be estimated ($\underline{\theta}$).

$$\underline{x}_i = \underline{\xi}_i + k\underline{\varepsilon}_i \text{ where } i = 1, 2\ldots, n \tag{10}$$

$$\underline{g}\left(\underline{\xi}_i, \underline{\theta}\right) = 0 \text{ where } i = 1, 2\ldots, n \tag{11}$$

In both Equations (10) and (11), i represents the trial number (out of n trials), and the underlined terms are either vectors or matrices.

The goal of the EVM is to minimize the sum of squares between the observed and predicted values, both in terms of the parameter estimates and the process variables [24]. The nested-iterative EVM algorithm [25] accomplishes this by using two nested loops, where the outer loop searches

for parameter estimates while the inner loop identifies estimates of the true values of the variables involved. Mathematically, the following objective function (Equation (12)) is minimized:

$$\phi = \frac{1}{2} \sum_{i=1}^{n} r_i (\overline{x}_i - \hat{\underline{\xi}}_i)' \underline{V}^{-1} (\overline{x}_i - \hat{\underline{\xi}}_i) \tag{12}$$

where r_i is the number of replicates for the ith trial, $\overline{x}_i$ is the average of the r_i measurements $(\underline{x}_i)$, $\hat{\underline{\xi}}_i$ is an estimate of the true values of the variables $(\underline{\xi}_i)$, and $\underline{V}$ is the variance-covariance matrix of the variables (which provides information about measurement error of the variables involved).

In addition to estimating model parameters and model variables, the EVM also evaluates the precision of the parameter estimates. Since several parameters are being estimated simultaneously, JCRs are plotted. JCRs are typically elliptical contours that quantify the level of uncertainty in the parameter estimates. Smaller JCRs indicate higher precision and therefore more confidence in the estimation results.

3. Results

In what follows, two main challenges associated with parameter estimation will be considered: Composition limitations (process constraints) and ill-conditioned systems (numerical estimation constraints). The discussion will be primarily informed by case studies from the literature. Then, we will revisit additional case studies that have shown good model agreement when using binary reactivity ratios to predict terpolymer properties. This provides a comparison of binary and ternary prediction performance for a variety of terpolymerization systems.

3.1. Addressing Composition Restrictions & Ill-Conditioned Systems

In some cases, system limitations and/or product requirements do not necessarily allow for statistically designed experiments, such as those described in Section 2 (each rich in a single comonomer). This paucity of data presents a challenge for researchers, especially when they hope to estimate reactivity ratios for subsequent microstructural predictions. The question is: What do we do when a terpolymer system has composition restrictions? How can we analyze a limited dataset in such a way that we are confident in our parameter estimates? In what follows, we take a closer look at the challenges associated with two such systems and provide some suggestions for overcoming the experimental limitations.

3.1.1. HOST/EAMA/PAG

In a recent study, Pujari et al. [5] synthesized a terpolymer of 4-hydroxystyrene (HOST; monomer 1), 2-ethyl-2-adamantyl methacrylate (EAMA; monomer 2), and a photoacid generator (PAG; monomer 3) to use in chemically amplified resists. Specifically, the PAG used for terpolymerization was triphenylsulfonium salt 4-(methacryloxy)-2,3,5,6-tetrafluorobenzenesulfonate (F4 PAG).

Pujari et al. [5] provided an interesting contrast between the copolymerization and terpolymerization behavior for the same comonomers. The authors reported that of the three comonomers, only HOST would homopolymerize. The HOST/EAMA copolymers were easily synthesized and HOST/PAG copolymers were achievable when the feed composition of PAG was below 10 mol%. However, their attempts to synthesize an EAMA/PAG copolymer were entirely unsuccessful (at a variety of distinct feed compositions). Finally, they synthesized HOST/EAMA/PAG terpolymers from several different feed compositions, in spite of their inability to synthesize some of the analogous copolymers. This is an example of a system where binary (copolymerization) data would not be suitable for predicting terpolymerization behavior.

Since Pujari et al. [5] had a specific application in mind, all terpolymer formulations had similar initial compositions. In the feed, mole fractions were selected within the following ranges: $0.25 \leq f_{1,0} \leq 0.40$, $0.50 \leq f_{2,0} \leq 0.75$, and $0.01 \leq f_{3,0} \leq 0.10$ (with $\sum f_i = 1.0$). Of particular interest for the current investigation is the very low mole fraction of PAG. At most, the PAG content in the feed was only 10 mol%. These experiments are not designed with reactivity ratio estimation in mind, rather,

application requirements take precedence. However, these (low conversion) terpolymerization data were still used to estimate reactivity ratios and predict terpolymer properties.

Pujari et al. [5] were able to estimate 'ternary' reactivity ratios using Procop 2.3 [26]. However, because this dataset has minimal information content, the results exhibit a multiplicity of solutions. Table 1 compares the reactivity ratio estimates reported by Pujari et al. [5] with three successive estimations (labelled A through C) using the instantaneous model via the EVM (with the preliminary estimate that $r_{ij} = 0.500$ for all parameters).

Table 1. Reactivity ratio estimates for the terpolymerization of 4-hydroxystyrene (HOST)/ 2-ethyl-2-adamantyl methacrylate (EAMA)/photoacid generator (PAG). Data from [5].

	r_{12}	r_{21}	r_{13}	r_{31}	r_{23}	r_{32}
Reported by Pujari et al. [5]	0.05	0.12	0.81	0.12	0.05	0.79
Current Study (Error-in-Variables Model (EVM) Estimation)						
Inst. Estimation (A)	0.1063	0.2155	0.6754	0.0001	0.0877	0.0307
Inst. Estimation (B)	0.1767	0.2394	**99.9721**	0.0002	0.0897	0.7035
Inst. Estimation (C)	0.0003	0.1753	2.5329	0.0001	0.0728	**99.9997**

Clearly, the estimation stage and results using the instantaneous model with EVM are numerically unstable (and therefore unreliable). There are multiple solutions for this estimation. This is a limitation of using a dataset that only contains limited process information arising from PAG-poor formulations. Although common sense suggests that a reactivity ratio pair of $(r_{13}, r_{31}) = (99.9721, 0.0002)$ seems unlikely, as seen in estimation (B), it is numerically possible. This is a numerical artefact, as reactivity ratios reach the 'upper bound' (UB) of the parameter estimation program (UB = 100 for all parameter values during estimation).

All reactivity ratios presented in Table 1 give almost identical prediction performance when the PAG fraction is low. As shown in Figure 2, all model predictions 'fit' the experimental data (reported at one conversion level) equally well. The model predictions using the original reactivity ratio estimates reported by Pujari et al. [5] are not shown here, but they also 'fit' the experimental data well and fall within the ranges of Figure 2.
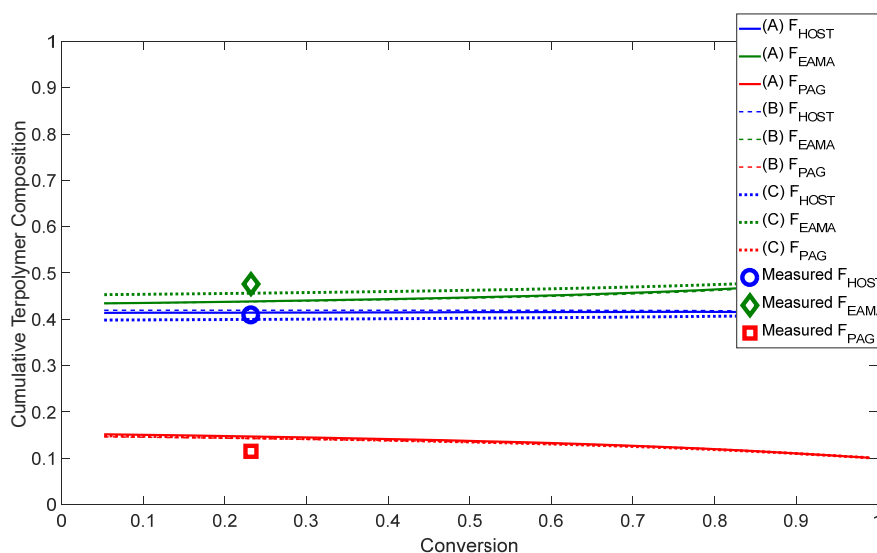

Figure 2. HOST/EAMA/PAG terpolymer composition prediction for $f_{1,0}/f_{2,0}/f_{3,0} = 0.4/0.5/0.1$ (experimentally measured composition for TPF10 from Pujari et al. [5]).

In contrast, if we predict terpolymerization behavior beyond the available experimental data (e.g., a PAG-rich formulation: $f_{1,0}/f_{2,0}/f_{3,0} = 0.1/0.1/0.8$), these considerably different reactivity ratio estimates

give distinctly different results (see Figure 3). This is an extreme case, selected for demonstration purposes, and may not be achievable experimentally. However, this inconsistent prediction performance is observed for any PAG-rich recipe (with $f_{3,0}$ as low as 0.4).
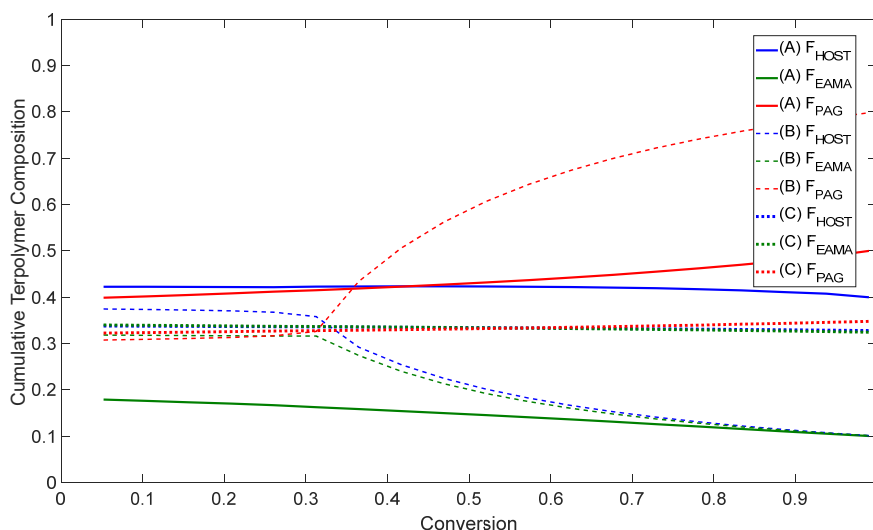

Figure 3. HOST/EAMA/PAG terpolymer composition prediction for $f_{1,0}/f_{2,0}/f_{3,0} = 0.1/0.1/0.8$.

The predictions shown in Figure 3 highlight the importance of statistically well-designed experiments. In spite of the experimental limitations in this case (limiting the PAG content to 10 mol%), introducing a run with as much PAG as possible would likely eliminate the numerical instabilities (and therefore improve the reliability of these estimation results).

3.1.2. BA/BMA/Limonene

Another terpolymerization case that is subject to composition restrictions (as well as ill-conditioning) is the terpolymerization of n-butyl acrylate (BA; monomer 1), butyl methacrylate (BMA; monomer 2), and D-limonene (lim; monomer 3), recently studied by Ren et al. [27]. D-limonene, a renewable monoterpene, is advantageous in terms of its polymer sustainability. When used as a comonomer it reduces the polymerization rate and molecular weight averages. Therefore, no more than 40 mol% lim was used in any feed composition for the study. The feed compositions used in the original investigation [27] are shown in Figure 4.
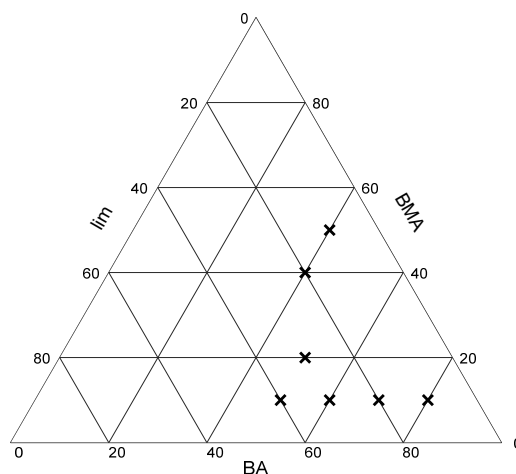

Figure 4. Feed compositions for the n-butyl acrylate (BA)/butyl methacrylate (BMA)/D-limonene (lim) terpolymer, as reported by Ren et al. [27].

Given that most experimental data are collected under lim-poor conditions, we might expect more error in the reactivity ratio estimates associated with the limonene comonomer. For both the instantaneous model (using low conversion data) and the cumulative model (using all available data), 'ternary' reactivity ratio estimation was performed three times (always with M_1 = BA, M_2 = BMA, and M_3 = lim). The results of each estimation are shown in Table 2. In this case, the binary reactivity ratio estimates (as reported by Ren et al. [27], collected from previous work by Dubé and colleagues [28–30]) were used as the preliminary estimates.

Table 2. Reactivity ratio estimates for the terpolymerization of BA/BMA/lim with experimental data from Ren et al. [27].

	r_{12}	r_{21}	r_{13}	r_{31}	r_{23}	r_{32}
Reported by Ren et al. [27]	0.46	2.008	6.08	0.007	6.096	0.046
Current Study (EVM Estimation with Low Conversion Data)						
Inst. Estimation (A)	0.3729	1.4350	5.0098	$<10^{-3}$	**35.5943**	**13.0409**
Inst. Estimation (B)	0.4986	1.4982	**5.2291**	**27.9826**	52.3811	$<10^{-3}$
Inst. Estimation (C)	0.3729	1.4350	5.0098	$<10^{-3}$	**35.5924**	**36.9594**
Current Study (EVM Estimation with Full Conversion Data)						
Cum. Estimation (A)	0.2787	0.9949	4.9888	0.0001	**29.6738**	**12.2200**
Cum. Estimation (B)	0.2733	0.9214	5.0104	0.0039	**27.1517**	**10.0072**
Cum. Estimation (C)	0.4081	0.9987	**5.4698**	**2.4651**	30.0914	0.0001

As we saw for the HOST/EAMA/PAG system, the estimation is numerically unstable. However, in general, the cumulative estimation results seem less ill-conditioned than in the instantaneous analysis. This is likely due to the increased information content provided when analyzing composition data over the full conversion range with a cumulative model [31].

Looking closer at the instantaneous estimation results, we see relatively good agreement between instantaneous estimations A and C. However, in both cases, the r_{23} and r_{32} estimates are both much greater than 1. Such a case has not been observed in free-radical copolymerization. There are some reports in the literature that have shown both estimates >1, but this is likely due to experimental error or a different copolymerization model being active. In this case, it is likely due to error, but the degradative chain transfer mechanism (due to the presence of limonene) may also contribute here. The uncertainty in this system is confirmed by comparing instantaneous estimations A and C to instantaneous estimation B. The fact that estimation results based on the same dataset have considerable convergence issues (likely due to local optima) suggests that there is not sufficient information for reactivity ratio estimation directly from the terpolymerization data.

Similar behavior is observed for the cumulative case. Here, cumulative estimations A and B are similar, whereas the third estimation, estimation (C), shows more variation. Again, the estimation results indicate that two reactivity ratios (for a given comonomer pair) are both greater than 1. The same comonomer pairs are of concern here: The BMA/lim comonomer pair (r_{23} and r_{32}) in cumulative estimations A and B, and the BA/lim pair (r_{13} and r_{31}) in cumulative estimation C. This behavior can be examined further by plotting the joint confidence regions (JCRs, or error ellipses) for each of the point estimates obtained using cumulative analysis (Figure 5).

As shown in Figure 5a, the JCRs associated with r_{23} and r_{32} are very large. The largest uncertainty was for estimation A, but subsequent estimations showed similar results. In comparison, the JCRs for the other parameters look like point estimates. This indicates substantial and disproportionate uncertainty in the estimates, especially for the BMA/lim comonomer pair.

As we focus in on the other parameter estimates (for the BA/BMA pair and the BA/lim pair), the JCRs become much smaller (note the change in scale between Figure 5a,b). Comparing BA/BMA to BA/lim, the most uncertainty is clearly for the BA/lim comonomer pair (especially for estimation (C)). In comparison, the JCRs for BA/BMA are very small, which gives us a much higher degree of certainty compared to the other estimates (see Figure 5c).
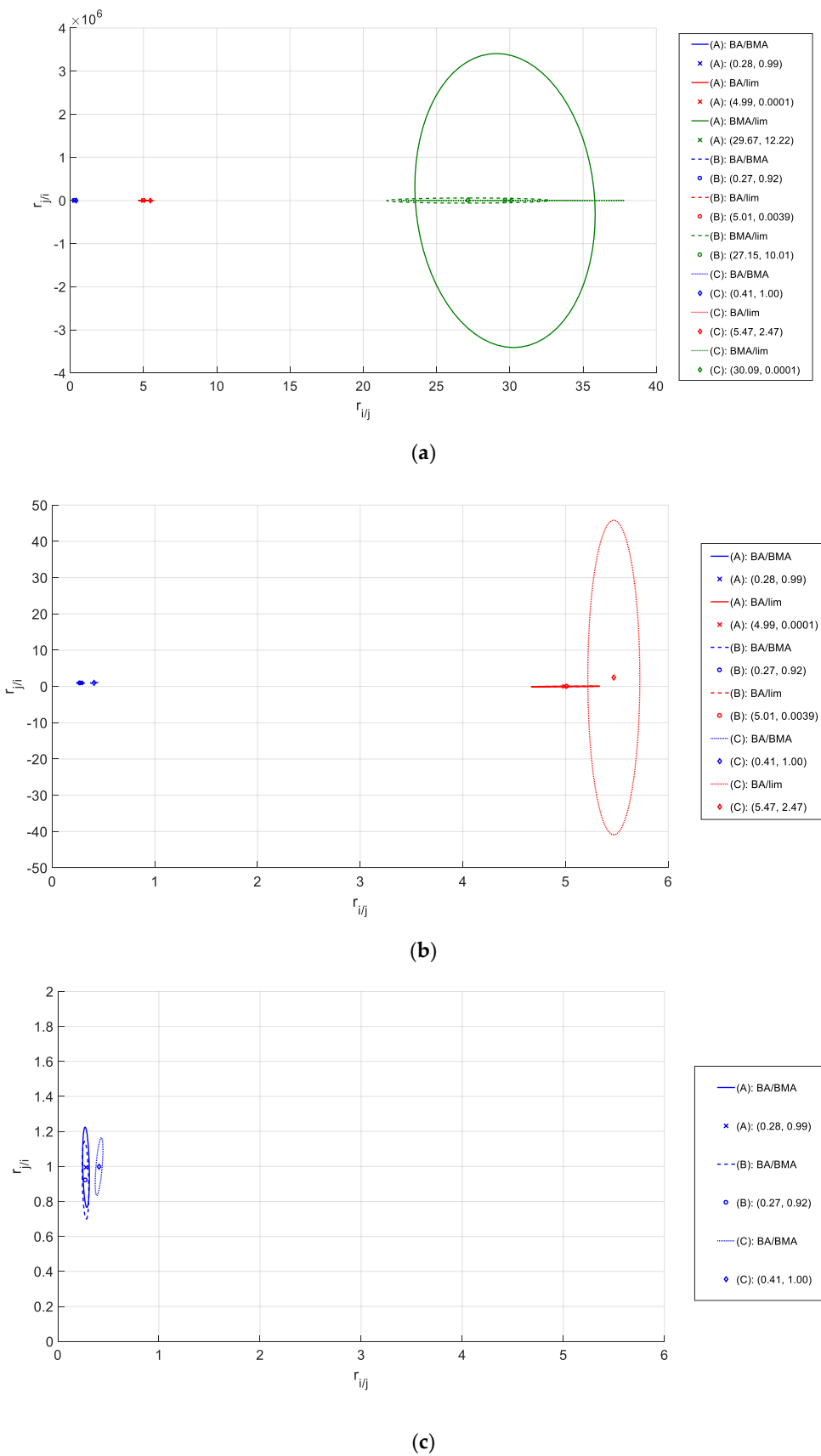
(**a**)

(**b**)

(**c**)

Figure 5. 'Ternary' reactivity ratio estimates for the terpolymerization of BA/BMA/lim, with data from Ren et al. [27].

The more precise estimation of r₁₂ and r₂₁ (that is, for the BA/BMA pair) is not coincidental. As mentioned earlier, the experimental data collected (Figure 4) only included formulations with low mole fractions of limonene. Therefore, we would suggest that a lack of lim-rich data has contributed to the poor estimation performance for the BA/BMA/lim terpolymer. This agrees with previous copolymerization observations reported by Scott et al. [31].

To demonstrate the importance of using well-designed data for 'ternary' reactivity ratio estimation, we have simulated supplemental data for the BA/BMA/lim system. Experimental data were simulated using the binary reactivity ratio estimates (which, as per the original investigation, give acceptable predictions of terpolymer behavior up to full conversion levels [27]). Two feed compositions, $f_{1,0}/f_{2,0}/f_{3,0}$ = 0.1/0.8/0.1 and $f_{1,0}/f_{2,0}/f_{3,0}$ = 0.1/0.1/0.8 (that is, BMA-rich and lim-rich formulations), were selected to supplement the original dataset. In both cases, the total conversion range was divided into 19 points (between 0 and 0.99, in steps of 0.052) and the corresponding monomer composition mole fractions were calculated via direct numerical integration. Then, the cumulative terpolymer compositions were calculated using the Skeist equation ([32], as per [22]). Random error was added to all data to mimic real experimental observations: A 1% error was added to the conversion and feed composition ($f_{i,0}$) data, while a 2% error was added to the cumulative terpolymer composition ($\bar{F}_i$) data. (Note that typically a 5% error is standard for $\bar{F}_i$ data, but using those levels for this system would occasionally make the simulated limonene content negative, given the low incorporation of limonene). The simulated data are provided in Table A1 of Appendix A.

For the instantaneous case, three low conversion data points from each (simulated) feed composition were added to the dataset. As shown in Table 3, the addition of these data 'stabilized' the estimation, and the reactivity ratio estimates obtained were now almost well-behaved. However, the newly estimated parameters still pose a concern: For the r_{23} and r_{32} pair (BMA/lim comonomers), both reactivity ratios were much greater than 1 for the low conversion case (instantaneous). In fact, r_{32} was estimated at 100.00 from three consecutive assessments. As explained previously (in Section 3.1.1), this is a numerical artefact, where the parameter has reached the 'upper bound' of the estimation program. However, the consistency of this result invites further investigation.

Table 3. Reactivity ratio estimates for terpolymerization of BA/BMA/lim with supplemental data (experimental data from [27] and simulated data from current work).

	r_{12}	r_{21}	r_{13}	r_{31}	r_{23}	r_{32}
Reported by Ren et al. [27]	0.46	2.008	6.08	0.007	6.096	0.046
Current Study (EVM Estimation with Experimental and Simulated Low Conversion Data)						
Inst. Estimation (A)	0.3663	1.4317	7.5324	0.1003	**10.1289**	**100.00**
Inst. Estimation (B)	0.3662	1.4317	7.5293	0.1004	**10.1307**	**100.00**
Inst. Estimation (C)	0.3663	1.4317	7.5324	0.1003	**10.1289**	**100.00**
Current Study (EVM Estimation with Experimental and Simulated Full Conversion Data)						
Cum. Estimation (A)	0.3009	2.4961	6.5236	0.0009	8.1873	0.0106
Cum. Estimation (B)	0.3167	1.9085	5.7962	0.0017	5.9647	0.0087
Cum. Estimation (C)	0.2876	1.5613	8.5751	0.0315	7.7202	0.1240

It is our understanding that the r_{32} = 100.00 result is due to the ill-conditioned nature of the terpolymer system. Physically, we can explain it as follows: The limonene incorporation is very low at low conversion levels, such that k_{32} (the rate constant for terminal limonene radicals adding BMA monomer units) is tending to 0, given the reactivity ratio definition, $r_{32} = k_{33}/k_{32}$, and as $k_{32} \to 0$, $r_{32} \to \infty$. With this logic, we can explain the observation that r_{32} is continually hitting the upper bound of the estimation program.

In contrast, the cumulative analysis (which uses the composition data, both experimental and simulated, over the full conversion range) gives more stable results. Specifically, the estimation results show that $r_{32} < 1$. Clearly, supplementing the terpolymerization dataset with optimal formulations (as per the EVM framework) significantly improves the stability and trustworthiness of ternary

parameter estimates. Ideally, even more experimental data would be collected for the lim-rich system to offset the error associated with low limonene incorporation.

3.2. Improved Performance with Ternary Data

The above examples present a variety of challenges: Datasets with low information content or ill-conditioned systems (with reactivity ratios of different orders of magnitude) can make estimation difficult. If the estimation steps are unstable, then one does not have confidence in the final estimates. Unfortunately, even with many data points, we do not always have all the required information. Therefore, the design of experiments for reactivity ratio estimation is key.

In the cases that follow, we present some case studies that highlight the advantages of analyzing terpolymerization data directly using 'ternary' reactivity ratio estimation. All three case studies were originally modelled using analogous binary reactivity ratios, where the terpolymerization data have been revisited and reanalyzed.

3.2.1. BA/MMA/EHA

A recent study by Gabriel and Dubé [12] investigated the terpolymer of BA (monomer 1), methyl methacrylate (MMA; monomer 2), and 2-ethylhexyl acrylate (EHA; monomer 3), which is a material of interest for pressure sensitive adhesives. First, the authors determined the reactivity ratio pairs for two of the associated copolymers (BA/EHA and MMA/EHA), and subsequently used these binary reactivity ratios, along with literature values for the BA/MMA reactivity ratios, to predict the terpolymer composition. The terpolymer model prediction (using binary reactivity ratios) showed good agreement with the collected data, as described in the original work [12].

In spite of the good results achieved using binary reactivity ratios, 'ternary' reactivity ratio estimation directly from terpolymerization data presents some additional advantages. First, we can consider the experimental load: Rather than nine experimental runs, as described by Gabriel and Dubé [12] (and additional prior work for estimating the BA/MMA reactivity ratios [33]), only three different feed compositions are required. Since Gabriel and Dubé [12] selected ternary feed compositions according to the EVM 'rule-of-thumb' for 'ternary' reactivity ratio estimation [21], we can use their data to re-estimate reactivity ratios directly from the terpolymerization data.

First, only the low conversion data were used for an instantaneous analysis. Now, because these data points were collected for model validation (not necessarily parameter estimation), only 7 data points are available below 20% conversion. These data (shown in Table A2 of the Appendix A) were used for 'ternary' reactivity ratio estimation using the recast Alfrey–Goldfinger equation (recall Equations (1)–(3) and the EVM (Equations (10)–(12)). There are two observations of note here: (1) The estimation is stable, much more so than the case studies presented in Section 3.1, and (2) the estimation is symmetrical. That is, regardless of which monomer is defined as monomer 1, monomer 2, or monomer 3, the estimated parameters are the same. As an example, two variations are shown below and compared to the original (binary) estimation. Here, reactivity ratios are labelled according to the monomer name (rather than monomer number) for further clarity. Also, the colors shown in Table 4 are associated with the colors of the JCRs in Figure 6.

Table 4. Reactivity ratio estimates for terpolymerization of BA/MMA/2-ethylhexyl acrylate (EHA) from low conversion data (experimental data from [12]).

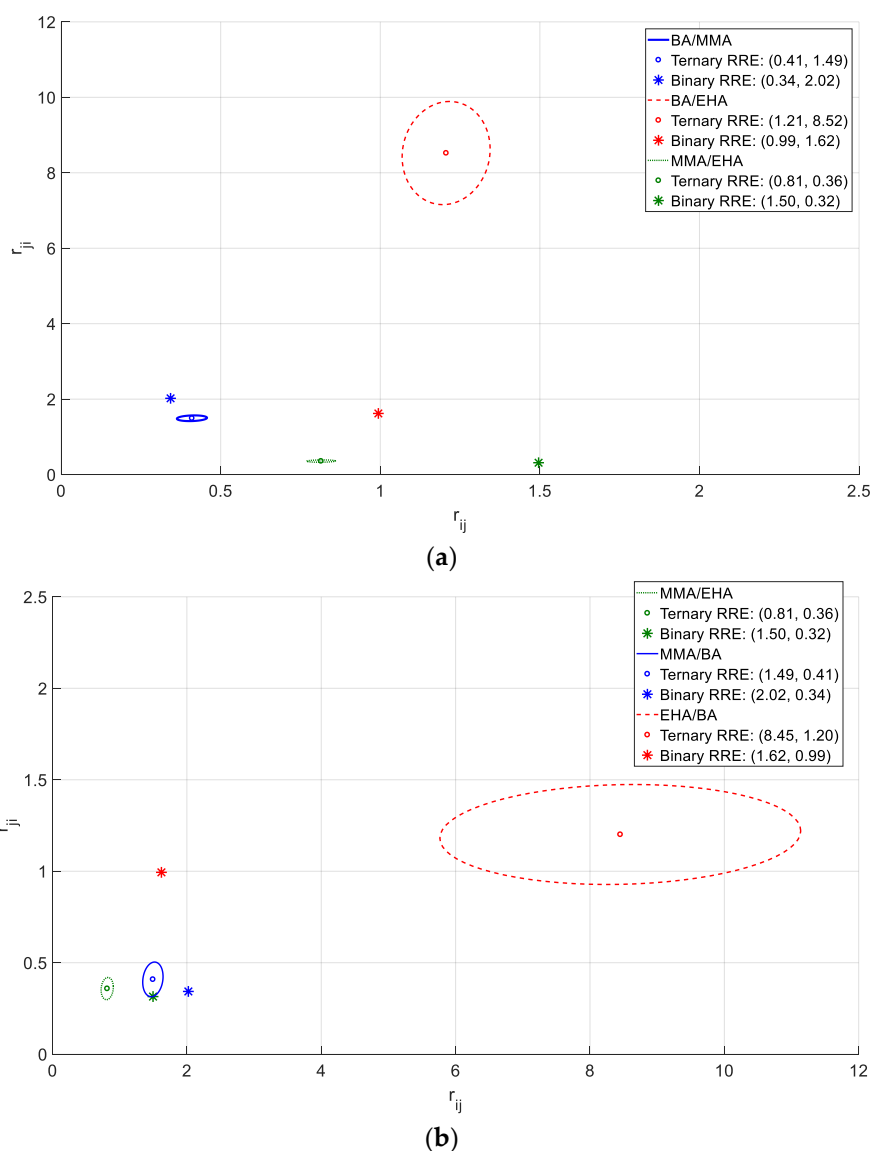|	$r_{BA/MMA}$	$r_{MMA/BA}$	$r_{BA/EHA}$	$r_{EHA/BA}$	$r_{MMA/EHA}$	$r_{EHA/MMA}$
Reported by Gabriel and Dubé [12] (binary reactivity ratio estimates)	0.34	2.02	0.99	1.62	1.50	0.32
Inst. Estimation ($M_1/M_2/M_3$ = BA/MMA/EHA)	0.41	1.49	1.21	8.52	0.81	0.36
Inst. Estimation ($M_1/M_2/M_3$ = MMA/EHA/BA)	0.41	1.49	1.20	8.45	0.81	0.36

Figure 6. 'Ternary' reactivity ratio estimates for the terpolymerization of BA/MMA/EHA for (**a**) $M_1/M_2/M_3$ = BA/MMA/EHA and (**b**) $M_1/M_2/M_3$ = MMA/EHA/BA, with instantaneous data from [12].

In general, the 'ternary' reactivity ratios follow the same trends as the original (binary) reactivity ratio estimates (that is, if $r_{ij} > r_{ji}$ for the binary case, the same relationship holds for the ternary case). However, $r_{MMA/EHA}$ falls below 1.00 when estimated directly from the terpolymerization data. This suggests that $k_{MMA/MMA} > k_{MMA/EHA}$ in the binary case (homopropagation of MMA is preferable to the crosspropagation of MMA and EHA), but that $k_{MMA/MMA} < k_{MMA/EHA}$ in the ternary case (homopropagation becomes dominated by the crosspropagation of MMA/EHA).

Another notable difference is the significant error present in the BA/EHA system. This is shown in both plots (of Figure 6), as the JCR for BA/EHA is much larger than the other JCRs. This may be related to the absolute value of the parameter estimates. As shown in a recent study [31], uncertainty becomes much greater for larger parameter values. Since $r_{EHA/BA}$ is larger than the other reactivity ratio estimates (by as much as 20 times, in some cases), the same relative error (assumed to be 5% for this system) will have a much larger absolute value in $r_{EHA/BA}$ compared to the other parameter estimates. This behavior has been described for the copolymer case [31], but the difference in parameter estimates was observed within a single JCR (that is, the elliptical JCR was stretched in the direction of

the larger parameter estimate). In this, a terpolymer case, the JCR associated with the comonomer pair containing larger parameter estimates is greater in both directions. The absolute value of the error seems magnified, likely because most other reactivity ratio estimates for the system are around or below 1.00. Another item of note is that both reactivity ratios for the BA/EHA pair are again >1. This may be for the same reasons discussed earlier for the BA/BMA/lim system. An additional reason may be related to the fact that the copolymerization of BA and EHA may lead to branched molecule formation and even microgel formation, which would complicate analysis further.

Next, we can look at the full conversion dataset, where 'ternary' reactivity ratios can be estimated using the EVM and the cumulative terpolymerization model. All terpolymerization data from the original study [12] were used herein, and the results are shown in Table 5 (and Figure 7). Again, the estimation is stable and symmetrical, which can be attributed to carefully designed data. As an aside, the estimation program also converged much more quickly, where parameters were estimated in under an hour (on an Intel(R) Core™ i7-860 processor) compared to (on average) 50 h of computation for the ill-conditioned system described earlier.

Table 5. Reactivity ratio estimates for the terpolymerization of BA/MMA/EHA from all terpolymerization data (experimental data from [12]).

	$r_{BA/MMA}$	$r_{MMA/BA}$	$r_{BA/EHA}$	$r_{EHA/BA}$	$r_{MMA/EHA}$	$r_{EHA/MMA}$
Reported by Gabriel and Dubé [12] (binary reactivity ratio estimates)	0.34	2.02	0.99	1.62	1.50	0.32
Cum. Estimation ($M_1/M_2/M_3$ = BA/MMA/EHA)	0.41	1.60	2.01	7.59	0.74	0.35
Cum. Estimation ($M_1/M_2/M_3$ = MMA/EHA/BA)	0.41	1.60	2.06	7.66	0.74	0.35

The values estimated using all terpolymerization data (full conversion) are similar to the results of the instantaneous parameter estimation (compare Table 4 to Table 5). Also, in comparing Figure 6 to Figure 7, the JCR areas are reduced when the full conversion dataset is used for analysis (note that the scales are the same for the easy comparison of Figure 6a to Figure 7a and of Figure 6b to Figure 7b). This is in agreement with previous studies within our group [18,22] and makes sense physically. Since more experimental data are available for analysis (18 data points over all conversion levels instead of 7 low conversion data points), the uncertainty associated with the parameter estimates is reduced. Also, since the instantaneous analysis used low conversion data up to 20%, the requisite assumption that no composition drift occurs may not be valid for all of the data [22]. Interestingly, a direct comparison of Figures 6 and 7 shows little or no JCR overlap (between the instantaneous and cumulative analysis), in spite of the fact that the trends remain consistent. Additional replication or sequential design of experiments could be used to further supplement this dataset, as has been described for the previous case studies (recall Section 3.1).

As for the instantaneous case, the most error associated with the cumulative analysis is present in the BA/EHA comonomer pair, and both reactivity ratios are greater than 1 (which is physically unlikely for any given comonomer pair). Again, this is likely due to the large values of the parameter estimates, which translate to a higher absolute value of the error (since we assume the same relative error for all experimental data and resulting parameter estimates). Also, as mentioned in the evaluation of the instantaneous results, other copolymerization mechanisms (branching, etc.) may be active specifically for the BA/EHA comonomer pair. However, without further analysis, no specific conclusions can be drawn about this system.

Finally, we can look at the prediction performance of these reactivity ratios (compared to the original binary reactivity ratio estimates). Since some of the more substantial differences in reactivity ratios were related to the EHA monomer (especially $r_{EHA/BA}$ and $r_{MMA/EHA}$), we can look at the model prediction for the EHA-rich terpolymer. The model predictions (using both 'binary' and 'ternary' reactivity ratios) and a comparison to the experimental data from Gabriel and Dubé [12] are shown

in Figure 8. Only the prediction performance of the cumulative analysis (for 'ternary' reactivity ratio estimation) is provided in Figure 8, where, despite slight differences between the instantaneous and cumulative analysis results, the model prediction performance was very similar for both sets of reactivity ratio estimates.
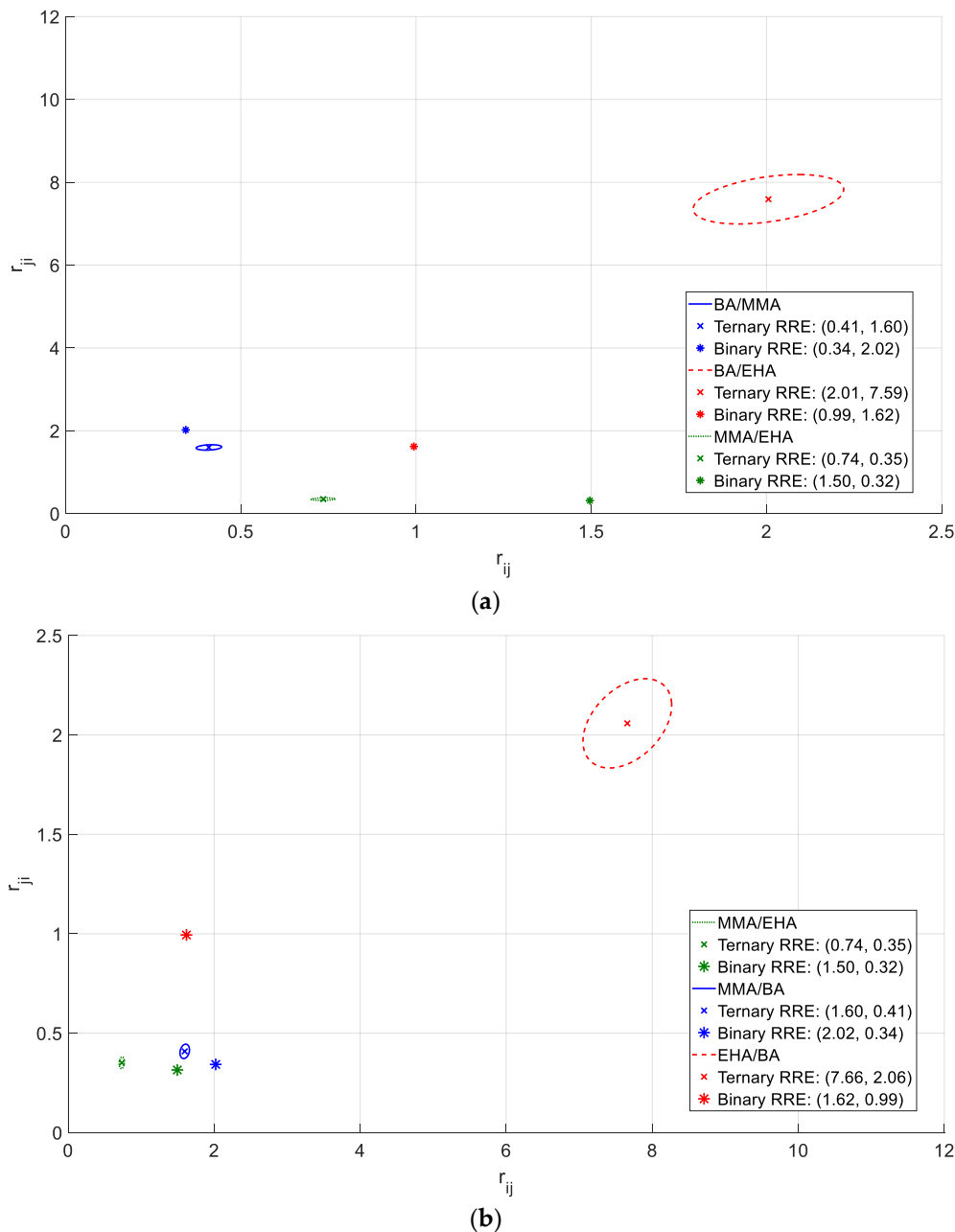

Figure 7. 'Ternary' reactivity ratio estimates for the terpolymerization of BA/MMA/EHA for (**a**) $M_1/M_2/M_3$ = BA/MMA/EHA and (**b**) $M_1/M_2/M_3$ = MMA/EHA/BA, with cumulative data from [12].

A direct comparison reveals that while the binary predictions are acceptable, the 'ternary' reactivity ratios further improve the prediction performance of the cumulative terpolymer composition model. In fact, a statistical comparison of the EHA-rich data (experimental data versus the two model predictions) shows that using the 'ternary' reactivity ratio estimates in the model leads to an 85% reduction in prediction error (total sum of square errors). Similar results were observed for the other terpolymer formulations but are not shown herein for the sake of brevity. These differences in

prediction performance may further be accentuated if the estimated reactivity ratios are used in the sequence length part of the model.
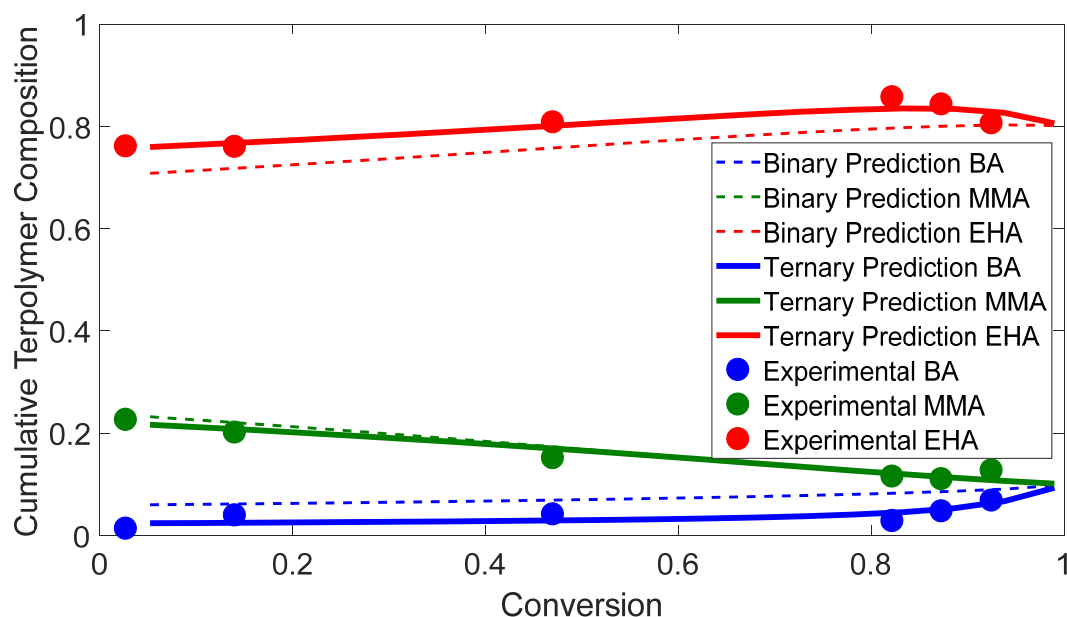

Figure 8. Prediction of cumulative terpolymer composition for BA/MMA/EHA ($f_{BA,0}/f_{MMA,0}/f_{EHA,0}$ = 0.1/0.1/0.8) (experimental data and binary predictions from [12]).

This case study has shown that when experiments are well-designed, 'ternary' reactivity ratio estimates can be obtained from small datasets. This allows for more resources to be directed towards careful replication and supplemental data collection. The results also suggest that binary and 'ternary' reactivity ratio estimates may be similar when the comonomers have similar structures and the polymerization is not affected by the solution properties. However, binary reactivity ratios are not always applicable to terpolymer systems (as has been shown recently [18,20]). In this case, the binary reactivity ratios gave reasonable prediction performance, but the 'ternary' reactivity ratios showed even better prediction performance based on fewer experimental data (and, hence, less effort overall).

3.2.2. Sty/MMA/MA

A study by Schoonbrood [34] looked at the terpolymerization kinetics for the styrene (Sty)/MMA/methyl acrylate (MA) (monomer 1/monomer 2/monomer 3 = Sty/MMA/MA) terpolymer. Again, according to standard practice, binary reactivity ratios (obtained from copolymerizations in the literature) were used to predict terpolymerization behavior. During this study, only low conversion data were reported. At the time (1994), this was 'best practice', where low conversion (instantaneous) data were typically used for parameter estimation. Parameter estimation from cumulative composition data was not part of typical practice, especially with the more complex system of equations representing terpolymerization kinetics. Low conversion data allow for a computationally simpler parameter estimation process but require some assumptions about a lack of composition drift in the system [22].

The experimentally determined (assumed as instantaneous) terpolymer compositions were compared to the model prediction. As reported in the original work, good agreement was observed between the predicted and measured values [34]. Given the available terpolymerization data, we can use the recast Alfrey–Goldfinger model (with the EVM) to re-estimate the terpolymer reactivity ratios directly from the terpolymerization data. The estimation is stable and symmetrical. A comparison of reactivity ratio estimates is presented in Table 6 and the prediction performance is evaluated in Table 7.

Table 6. Reactivity ratio estimates for the terpolymerization of styrene (Sty)/MMA/methyl acrylate (MA). Experimental data from [34].

	r_{12}	r_{21}	r_{13}	r_{31}	r_{23}	r_{32}
Reported by Schoonbrood [34] (binary reactivity ratio estimates)	0.48	0.42	0.73	0.19	2.49	0.29
Instantaneous Estimation (current work)	0.57	0.51	1.82	0.20	2.49	0.23

Table 7. Comparison of model predictions for the Sty/MMA/MA terpolymerization. Experimental data and original predictions from [34].

Feed Composition			Experimental Data			Original Predictions			Current (EVM) Predictions		
$f_{sty,0}$	$f_{MMA,0}$	$f_{MA,0}$	$\overline{F}_{sty}$	$\overline{F}_{MMA}$	$\overline{F}_{MA}$	$\overline{F}_{sty}$	$\overline{F}_{MMA}$	$\overline{F}_{MA}$	$\overline{F}_{sty}$	$\overline{F}_{MMA}$	$\overline{F}_{MA}$
0.10	0.10	0.80	0.27	0.28	0.45	0.26	0.19	0.55	0.26	0.23	0.51
0.10	0.20	0.70	0.24	0.38	0.38	0.24	0.33	0.44	0.23	0.37	0.40
0.10	0.30	0.60	0.21	0.45	0.33	0.23	0.43	0.34	0.21	0.48	0.31
0.20	0.10	0.70	0.44	0.18	0.39	0.37	0.16	0.47	0.40	0.20	0.41
0.20	0.20	0.60	0.37	0.39	0.24	0.35	0.28	0.37	0.36	0.33	0.31
0.20	0.30	0.50	0.33	0.39	0.28	0.34	0.38	0.28	0.34	0.43	0.23
0.20	0.50	0.30	0.29	0.60	0.11	0.32	0.53	0.15	0.31	0.57	0.12
0.30	0.20	0.50	0.42	0.27	0.31	0.43	0.26	0.31	0.46	0.31	0.24
0.30	0.30	0.40	0.40	0.39	0.20	0.42	0.35	0.23	0.43	0.40	0.17
0.30	0.40	0.30	0.41	0.44	0.15	0.40	0.43	0.16	0.41	0.47	0.12
0.30	0.50	0.20	0.39	0.54	0.07	0.39	0.50	0.10	0.39	0.54	0.07
0.40	0.30	0.30	0.48	0.34	0.18	0.48	0.34	0.18	0.50	0.38	0.12
0.40	0.40	0.20	0.50	0.36	0.14	0.47	0.42	0.11	0.47	0.45	0.08
0.50	0.20	0.30	0.56	0.24	0.21	0.55	0.25	0.21	0.59	0.28	0.13
0.50	0.40	0.10	0.50	0.49	0.01	0.52	0.42	0.06	0.53	0.43	0.04
0.50	0.30	0.20	0.52	0.43	0.06	0.53	0.34	0.13	0.56	0.36	0.08

Although the prediction performance looks similar, the current (EVM) prediction shows a decrease in the sum of square errors for all three comonomer compositions, especially $\overline{F}_{MMA}$ and $\overline{F}_{MA}$. In evaluating the total sum of square errors, the current work provides a 32% decrease in prediction error over the original analysis. To supplement this result, we can also examine the residuals for both the original and current predictions. As shown in Figure 9, the spread (that is, the vertical distance from 0) is reduced for the current predictions, where the residuals are smaller overall.
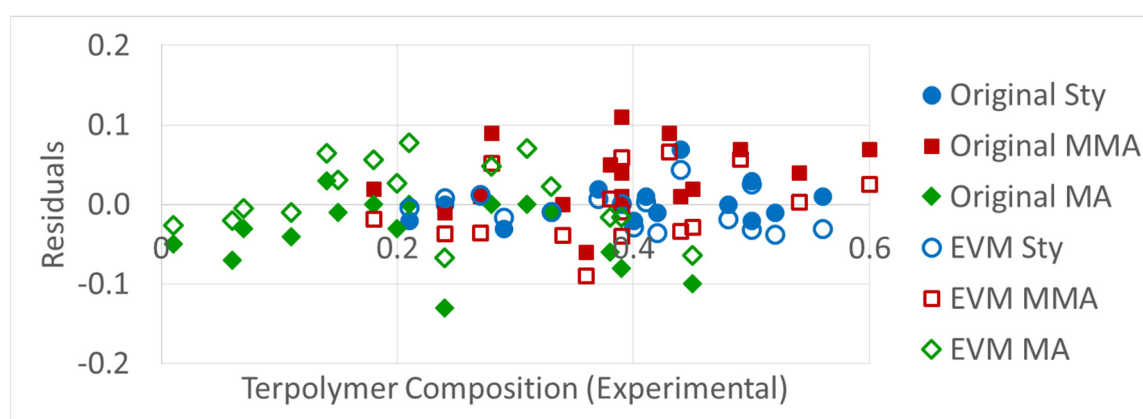
Figure 9. Comparison of residuals for Sty/MMA/MA terpolymer composition predictions.

Given these results (and those discussed previously), there is clearly an advantage for estimating 'ternary' reactivity ratios directly from terpolymerization data. If medium to high conversion data were available, they could have been used to supplement the dataset or to reduce the number of

experiments required. However, even with this low conversion dataset, estimating 'ternary' reactivity ratios directly from terpolymerization data is feasible and preferable to using binary data.

3.2.3. AN/Sty/MMA

The terpolymerization of acrylonitrile (AN; monomer 1), Sty (monomer 2), and MMA (monomer 3) studied by Brar and Hekmatyar [35] provides us with some interesting experimental data. In addition to reporting the terpolymer composition data, they also reported microstructural (triad fraction) information. Thus, there is potential to re-estimate the 'ternary' reactivity ratios for AN/Sty/MMA and evaluate their ability to predict composition and sequence length distribution.

The original investigation used six experiments (no replication is mentioned) and the feed compositions selected provide a good amount of experimental information. As shown in the triangular diagram of Figure 10, there are three 'outer' formulations further along the outside of the triangle (designated with circles in Figure 10). Although (to the best of our knowledge) these were not statistically designed experiments, the fact that there is one formulation rich in each comonomer provides useful data for reactivity ratio estimation [21]. In fact, when the reactivity ratios are estimated using only these three trials, the parameter estimation results are as expected.
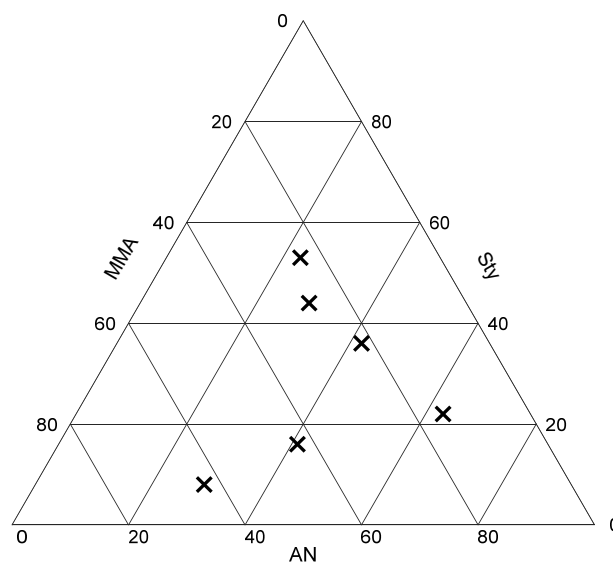

Figure 10. Terpolymerization feed compositions for AN/Sty/MMA. Data from [35].

Ideally, additional replication would be performed for all six formulations (perhaps even to higher conversion levels), but these results certainly represent carefully measured experimental (process) data (hence, good information content/lower experimental error), leading to a numerically stable estimation situation, thus ensuring that reactivity ratios can be successfully estimated even from a limited dataset.

Figure 11 shows a comparison of estimation results from these three (more optimal) points to the estimation results from the full (six) trial set. Clearly, the three 'internal' data points supplement the composition data, but do not significantly alter the reactivity ratio estimation results. Also, the 'ternary' reactivity ratio estimates are in good agreement with the previously reported binary reactivity ratios for the associated copolymers.

For all three comonomer pairs, the binary reactivity ratios are within the JCRs for the ternary estimates. Thus, the prediction performance (for both terpolymer composition and microstructure) will be similar, regardless of which parameters are used. For the purposes of demonstration, analysis of the composition and microstructure of one terpolymer sample (experimentally determined by Brar and Hekmatyar [35]) is summarized in Table 8. Note that triad fractions are defined only by the first letter of the monomer name, for example, triad fraction ASM represents the AN-Sty-MMA triad sequence.
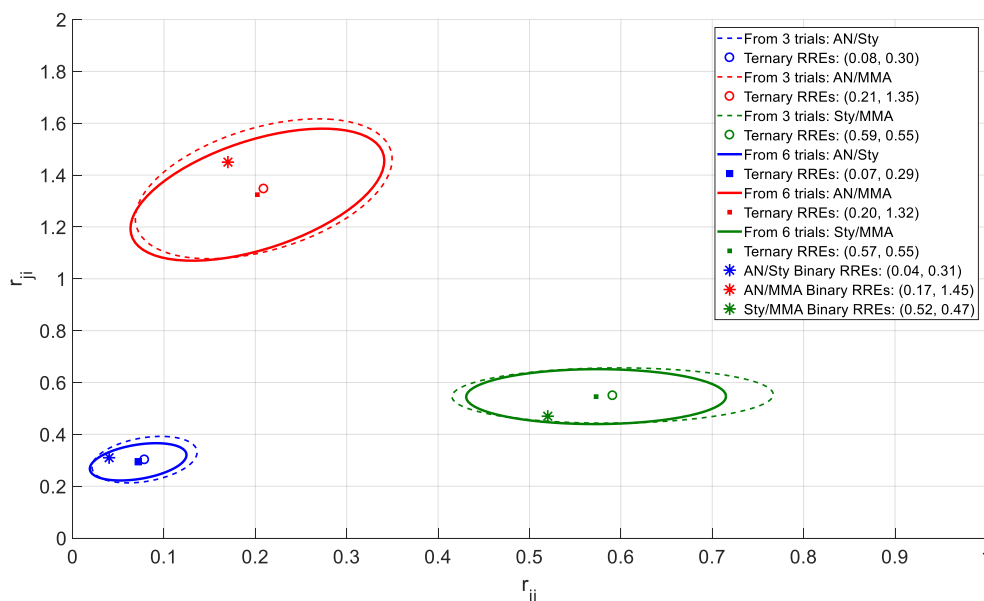
Figure 11. 'Ternary' reactivity ratio estimates (RREs) for the terpolymerization of AN/Sty/MMA with the data from Brar and Hekmatyar [35].

Table 8. Analysis of (**a**) composition and (**b**) microstructure for AN/Sty/MMA (experimental data from [35]). RRE = reactivity ratio estimates.

(a)			
Monomer	**Experimental [35]**	**Original prediction (from binary RREs) [35]**	**Current prediction (from ternary RREs)**
AN	0.30	0.28	0.30
Sty	0.48	0.48	0.48
MMA	0.22	0.24	0.22
(b)			
Triad	**Experimental [35]**	**Original prediction (from binary RREs) [35]**	**Current prediction (from ternary RREs)**
AAA	0.01	0.01	0.00
SAS	0.69	0.73	0.63
MAM	0.03	0.00	0.03
AAS	0.06	0.04	0.06
AAM	0.01	0.00	0.01
SAM	0.20	0.22	0.27
SSS	0.05	0.05	0.05
ASA	0.21	0.25	0.27
MSM	0.08	0.08	0.06
SSA	0.27	0.22	0.24
SSM	0.16	0.12	0.12
ASM	0.23	0.28	0.26
MMM	0.05	0.04	0.04
AMA	0.05	0.02	0.03
SMS	0.42	0.40	0.39
MMA	0.04	0.06	0.07
MMS	0.21	0.26	0.26
AMS	0.23	0.22	0.21

4. Conclusions

Through a series of case studies, we have demonstrated with examples and counter-examples both the challenges and advantages of estimating 'ternary' reactivity ratios directly from terpolymerization data. We highlighted some difficulties that may arise when studying multi-component polymerizations due to the nature of such systems and related experimental limitations. These limitations are usually translated to the paucity of experimental information content. The lack of sufficient information on polymer composition, combined with uncertain levels of experimental error (since independent replication is usually non-existent) usually leads to numerically ill-conditioned systems during the estimation steps. This in its turn results in a multiplicity of (and

related confusion with) reactivity ratio values. Now, if the above already important limitations are superimposed to a lack of experimental design (i.e., the optimal selection of feed compositions), the problem is compounded with the extra dimension of added pitfalls with the use of undesigned (happenstance) data.

When experimental design is employed from the outset and is combined with appropriate parameter estimation techniques using carefully measured (and replicated) data, i.e., when every effort is made to make the terpolymerization system numerically 'well-behaved', then accurate and reliable estimates of 'ternary' reactivity ratios can be obtained. The case studies examined gave examples of both instantaneous (low conversion) and cumulative (medium-high conversion) data analysis and demonstrated the advantages of using the cumulative model. Some systems exhibited similarities between the 'binary' and 'ternary' reactivity ratio estimates, and the predictions related to composition and triad fractions were improved.

In some other terpolymerization systems, the reactivity ratios of the binary copolymerization pairs do not apply to ternary systems [20]. Trying to predict terpolymerization behavior from binary reactivity ratios can require making unfounded assumptions about system-specific polymerization kinetics. Researchers often assume that the presence of the third comonomer is exactly additive via simple superposition, hence the behavior of the three comonomers is independent of each other. This is the main assumption, akin to assuming that interaction terms are non-existent in a model. Evaluating subsets of the experimental data collected (that is, comonomer pairs) to represent ternary systems can ultimately result in the oversimplification of complex processes. Using appropriate kinetic models can result in better prediction performance and a higher degree of confidence in the resulting parameter estimates.

Overall, and as we have demonstrated in the current work, 'ternary' reactivity ratio estimation directly from terpolymerization data can provide improved understanding about a complex terpolymer system. Dealing for the first time with a completely unknown terpolymerization system, the current approach offers, if nothing else, a systematic and 'safe' approach to accumulating experimental evidence about the system in question in fewer experimental trials, and hence with less experimental effort, but with reliable parameter estimates which can be fine-tuned further later, as one becomes more familiar with the terpolymerization system in a sequential-iterative-optimal manner.

Appendix A Additional Data for Case Studies

Table A1. Simulated data for the BA/BMA/lim terpolymerization (supplementing experimental data from Ren et al. [27]; see Section 3.1.2).

X	$f_{BA,0}$	$f_{BMA,0}$	$f_{lim,0}$	$\bar{F}_{BA}$	$\bar{F}_{BMA}$	$\bar{F}_{lim}$
*0.041	0.098	0.791	0.111	0.053	0.931	0.016
*0.102	0.101	0.793	0.106	0.086	0.904	0.010
*0.153	0.100	0.802	0.098	0.044	0.928	0.028
0.221	0.111	0.800	0.089	0.038	0.916	0.046
0.269	0.085	0.804	0.112	0.072	0.895	0.033
0.313	0.101	0.801	0.098	0.055	0.936	0.010
0.361	0.103	0.795	0.102	0.062	0.897	0.041
0.420	0.091	0.808	0.102	0.076	0.890	0.034

Table A1. *Cont.*

X	$f_{BA,0}$	$f_{BMA,0}$	$f_{lim,0}$	$\bar{F}_{BA}$	$\bar{F}_{BMA}$	$\bar{F}_{lim}$
0.460	0.102	0.797	0.101	0.066	0.904	0.031
0.514	0.109	0.801	0.090	0.054	0.894	0.052
0.578	0.114	0.794	0.092	0.039	0.931	0.030
0.620	0.097	0.802	0.101	0.089	0.893	0.018
0.671	0.108	0.799	0.093	0.090	0.879	0.031
0.728	0.094	0.801	0.106	0.119	0.866	0.015
0.780	0.100	0.802	0.097	0.062	0.893	0.046
0.833	0.102	0.798	0.100	0.072	0.880	0.047
0.892	0.107	0.797	0.095	0.052	0.881	0.067
0.936	0.117	0.797	0.085	0.122	0.838	0.041
0.983	0.109	0.793	0.098	0.119	0.795	0.086
*0.055	0.088	0.107	0.805	0.376	0.333	0.292
*0.110	0.103	0.093	0.805	0.387	0.294	0.319
*0.160	0.098	0.101	0.801	0.367	0.317	0.316
0.209	0.088	0.109	0.803	0.366	0.298	0.337
0.260	0.119	0.096	0.785	0.363	0.286	0.351
0.308	0.091	0.103	0.806	0.322	0.269	0.409
0.359	0.099	0.105	0.796	0.249	0.275	0.476
0.419	0.101	0.099	0.801	0.225	0.246	0.528
0.469	0.098	0.101	0.800	0.215	0.211	0.574
0.520	0.106	0.090	0.804	0.159	0.207	0.634
0.576	0.102	0.098	0.801	0.168	0.183	0.648
0.624	0.090	0.100	0.809	0.138	0.180	0.683
0.672	0.116	0.091	0.793	0.152	0.162	0.686
0.727	0.103	0.100	0.797	0.129	0.149	0.722
0.781	0.112	0.090	0.798	0.112	0.129	0.759
0.834	0.097	0.101	0.802	0.135	0.113	0.752
0.890	0.105	0.097	0.798	0.111	0.120	0.770
0.935	0.098	0.102	0.800	0.112	0.120	0.769
0.993	0.087	0.110	0.803	0.102	0.092	0.807

Table A2. Low conversion data for the BA/MMA/EHA terpolymerization (experimental data from Gabriel and Dubé [12]; see Section 3.2.1).

X	$f_{BA,0}$	$f_{MMA,0}$	$f_{EHA,0}$	$\bar{F}_{BA}$	$\bar{F}_{MMA}$	$\bar{F}_{EHA}$
0.027	0.098	0.097	0.805	0.013	0.226	0.761
0.140	0.098	0.097	0.805	0.039	0.201	0.760
0.033	0.100	0.800	0.100	0.039	0.843	0.118
0.088	0.100	0.800	0.100	0.062	0.824	0.114
0.171	0.100	0.800	0.100	0.087	0.814	0.099
0.003	0.800	0.099	0.100	0.654	0.236	0.110
0.026	0.800	0.099	0.100	0.654	0.236	0.110

References

1. Rudin, A. Copolymerization. In *Elements of Polymer Science and Engineering: An Introductory Text and Reference for Engineers and Chemists*; Elsevier Science & Technology Books: Amsterdam, The Netherlands, 1998; Volume 7, pp. 241–276.
2. Scott, A.J.; Penlidis, A. Copolymerization. In *Reference Module*; Elsevier: Amsterdam, The Netherlands, 2017.
3. Smith, O.L.; Kim, Y.; Kathaperumal, M.; Gadinski, M.R.; Pan, M.; Wang, Q.; Perry, J.W. Enhanced permittivity and energy density in neat poly (vinylidene fluoride-trifluoroethylene-chlorotrifluoroethylene) terpolymer films through control of morphology. *ACS Appl. Mater. Interfaces* **2014**, *6*, 9584–9589. [CrossRef] [PubMed]
4. Li, Q.; Wang, Q. Ferroelectric polymers and their energy-related applications. *Macromol. Chem. Phys.* **2016**, *217*, 1228–1244. [CrossRef]

5. Pujari, N.S.; Wang, M.; Gonsalves, K. Co and terpolymer reactivity ratios of chemically amplified resists. *Polymer* **2017**, *118*, 201–214. [CrossRef]

6. Zhou, J.; Liu, J.; Cheng, C.J.; Patel, T.R.; Weller, C.E.; Piepmeier, J.M.; Jiang, Z.; Saltzman, W.M. Biodegradable poly (amine-co-ester) terpolymers for targeted gene delivery. *Nat. Mater.* **2012**, *11*, 82–90. [CrossRef] [PubMed]

7. Chung, H.; Grubbs, R.H. Rapidly cross-linkable DOPA containing terpolymer adhesives and PEG-based cross-linkers for biomedical applications. *Macromolecules* **2012**, *45*, 9666–9673. [CrossRef]

8. Chauhan, N.P.S. Preparation and characterization of bio-based terpolymer derived from vanillin oxime, formaldehyde, and p-hydroxyacetophenone. *Des. Monomers Polym.* **2014**, *17*, 176–185. [CrossRef]

9. Mecham, S.; Sentman, A.; Sambasivam, M. Amphiphilic silicone copolymers for pressure sensitive adhesive applications. *J. Appl. Polym. Sci.* **2010**, *116*, 3265–3270. [CrossRef]

10. Bowman, H.; Hudson, B.W. Terpolymers as Pressure Sensitive Adhesives. U.S. Patent 9598532B2, 21 March 2017.

11. Roberge, S.; Dubé, M.A. Emulsion-based pressure sensitive adhesives from conjugated linoleic acid/styrene/butyl acrylate terpolymers. *Int. J. Adhes. Adhes.* **2016**, *70*, 17–25. [CrossRef]

12. Gabriel, V.A.; Dubé, M.A. Bulk free-radical co- and terpolymerization of n-butyl acrylate/2-ethylhexyl acrylate/methyl methacrylate. *Macromol. React. Eng.* **2018**, *13*, 1800057. [CrossRef]

13. Hong, F.; Xie, L.; He, C.; Liu, J.; Zhang, G.; Wu, C. Effects of hydrolyzable comonomer and cross-linking on anti-biofouling terpolymer coatings. *Polymer* **2013**, *54*, 2966–2972. [CrossRef]

14. Pollack, K.A.; Imbesi, P.M.; Raymond, J.E.; Wooley, K.L. Hyperbranched fluoropolymer-polydimethylsiloxane-poly(ethylene glycol) cross-linked terpolymer networks designed for marine and biomedical applications: Heterogeneous nontoxic antibiofouling surfaces. *ACS Appl. Mater. Interfaces* **2014**, *6*, 19265–19274. [CrossRef] [PubMed]

15. Yousefi, F.K.; Jannesari, A.; Pazokifard, S.; Saeb, M.R.; Scott, A.J.; Penlidis, A. Terpolymerization of triisopropylsilyl acrylate, methyl methacrylate, and butyl acrylate: Reactivity ratio estimation. *Macromol. React. Eng.* **2019**, 1900014. [CrossRef]

16. Slocombe, R. Multicomponent polymers. I. Three-component systems. *J. Polym. Sci.* **1957**, *XXVI*, 9–22. [CrossRef]

17. Hagiopol, C. *Copolymerization: Toward a Systematic Approach*; Plenum Publishers: New York, NY, USA, 1999.

18. Kazemi, N.; Duever, T.A.; Penlidis, A. Demystifying the estimation of reactivity ratios for terpolymerization systems. *AIChE J.* **2014**, *60*, 1752–1766. [CrossRef]

19. Scott, A.J.; Duever, T.A.; Penlidis, A. The role of pH, ionic strength and monomer concentration on the terpolymerization of 2-acrylamido-2-methylpropane sulfonic acid, acrylamide and acrylic acid. *Polymer* **2019**, *177*, 214–230. [CrossRef]

20. Scott, A.J.; Penlidis, A. Binary vs. ternary reactivity ratios: Appropriate estimation procedures with terpolymerization data. *Eur. Polym. J.* **2018**, *105*, 442–450. [CrossRef]

21. Kazemi, N.; Duever, T.A.; Penlidis, A. Design of optimal experiments for terpolymerization reactivity ratio estimation. *Macromol. React. Eng.* **2015**, *9*, 228–244. [CrossRef]

22. Kazemi, N.; Duever, T.A.; Penlidis, A. Reactivity ratio estimation from cumulative copolymer composition data. *Macromol. React. Eng.* **2011**, *5*, 385–403. [CrossRef]

23. Alfrey, T.; Goldfinger, G. Copolymerization of systems of three and more components. *J. Chem. Phys.* **1944**, *12*, 322–323. [CrossRef]

24. Kazemi, N.; Duever, T.A.; Penlidis, A. A powerful estimation scheme with the error-in-variables model for nonlinear cases: Reactivity ratio estimation examples. *Comput. Chem. Eng.* **2013**, *48*, 200–208. [CrossRef]

25. Reilly, P.M.; Reilly, H.V.; Keeler, S.E. Parameter estimation in the error-in-variables model. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1993**, *42*, 693–701.

26. Mao, R.; Huglin, M.B. A new linear method to calculate monomer reactivity ratios by using high conversion copolymerization data: Terminal model. *Polymer* **1993**, *34*, 1709–1715. [CrossRef]

27. Ren, S.; Zhang, L.; Dubé, M.A. Free-radical terpolymerization of n-butyl acrylate/butyl methacrylate/D-limonene. *J. Appl. Polym. Sci.* **2015**, *132*, 42821. [CrossRef]

28. Zhang, Y.; Dubé, M.A. Copolymerization of n-butyl methacrylate and D-limonene. *Macromol. React. Eng.* **2014**, *8*, 805–812. [CrossRef]

29. Ren, S.; Trevino, E.; Dubé, M.A. Copolymerization of limonene with n-butyl acrylate. *Macromol. React. Eng.* **2015**, *9*, 339–349. [CrossRef]

30. Ren, S.; Hinojosa-Castellanos, L.; Zhang, L.; Dubé, M.A. Bulk free-radical copolymerization of n-butyl acrylate and n-butyl methacrylate: Reactivity ratio estimation. *Macromol. React. Eng.* **2017**, *11*, 1600050. [CrossRef]

31. Scott, A.J.; Penlidis, A. Computational package for copolymerization reactivity ratio estimation: Improved access to the error-in-variables-model. *Processes* **2018**, *6*, 8. [CrossRef]

32. Skeist, I. Copolymerization: The composition distribution curve. *J. Am. Chem. Soc.* **1946**, *68*, 1781–1784. [CrossRef]

33. Dubé, M.A.; Penlidis, A. A systematic approach to the study of multicomponent polymerization kinetics—the butyl acrylate/methyl methacrylate/vinyl acetate example: 1. Bulk copolymerization. *Polymer* **1995**, *36*, 587–598. [CrossRef]

34. Schoonbrood, H.A.S. *Emulsion Co- and Terpolymerization: Monomer Partitioning, Kinetics and Control of Microstructure and Mechanical Properties*; Technische Universiteit Eindhoven: Eindhoven, The Netherlands, 1994.

35. Brar, A.S.; Hekmatyar, S.K. Microstructure determination of the acrylonitrile-styrene-methyl methacrylate terpolymers by NMR spectroscopy. *J. Appl. Polym. Sci.* **1999**, *74*, 3026–3032. [CrossRef]