A Graph Theoretical Approach for Testing Binomiality of Reversible Chemical Reaction Networks

Hamid Rahkooy
CNRS, Inria, and the University of Lorraine
Nancy, France
hamid.rahkooy@inria.fr

Cristian Vargas Montero
CNRS, Inria, and the University of Lorraine
Nancy, France
cristian.vargas-montero6@etu.univ-lorraine.fr

Abstract

We study binomiality of the steady state ideals of chemical reaction networks. Considering rate constants as indeterminates, the concept of unconditional binomiality has been introduced and an algorithm based on linear algebra has been proposed in a recent work for reversible chemical reaction networks, which has a polynomial time complexity upper bound on the number of species and reactions. In this article, using a modified version of species–reaction graphs, we present an algorithm based on graph theory which performs by adding and deleting edges and changing the labels of the edges in order to test unconditional binomiality. We have implemented our graph theoretical algorithm as well as the linear algebra one in Maple and made experiments on biochemical models. Our experiments show that the performance of the graph theoretical approach is similar to or better than the linear algebra approach, while it is drastically faster than Gröbner basis and quantifier elimination methods.

1 Introduction

Chemical reactions are transformations between chemical species where a creation or elimination of species may happen with respect to changes in time, pressure, temperature, etc. A set of chemical reactions is called a chemical reaction network (CRN) and if all the reactions in a CRN are reversible, it is called a reversible chemical reaction network (RCRN). We assume mass-action kinetics in this article. The following is an example of an RCRN.

$$A + B \xrightleftharpoons[k_{21}]{k_{12}} C \xrightleftharpoons[k_{32}]{k_{23}} A + 2D,$$

with species A, B, C and D and rate constants $k_{12}, k_{21}, k_{23}$ and $k_{32}$.

Ordinary differential equations (ODE) can be used to study the changes in the concentration of species of a CRN. The ODEs associated to the above CRN are

$$\dot{x}_1 = p_1, \quad p_1 = -k_{12}x_1x_2 + k_{21}x_3 + k_{23}x_3 - k_{32}x_1x_4$$

$$\dot{x}_2 = p_2, \quad p_2 = -k_{12}x_1x_2 + k_{21}x_3$$

$$\dot{x}_3 = p_3, \quad p_3 = k_{12}x_1x_2 - k_{21}x_3 - k_{23}x_3 + k_{32}x_1x_4$$

$$\dot{x}_4 = p_4, \quad p_4 = 2k_{23}x_3 - 2k_{32}x_1x_4$$

(1)

The ideal generated by $p_1, p_2, p_3, p_4$ is called the steady state ideal of the CRN. The real positive zeros of the above ideal are called the steady states. Finding steady states is a fundamental problem in CRN theory. For a thorough introduction to CRN theory, we refer to Feinberg’s Book [12].

A CRN is called binomial if its steady state ideal is binomial. Following the work in [26], in
this article we investigate binomiality over the ring \( \mathbb{Q}[k_{ij}, x_1, \ldots, x_n] \), which we call **unconditional binomiality**. Some authors have considered binomiality of CRNs over \( \mathbb{Q}(k_{ij})[x_1, \ldots, x_n] \) when \( k_{ij} \) are specialised in \( \mathbb{R} \), e.g., in [24], which we call **conditional binomiality**.

Binomial ideals and toric varieties are historic topics in thermodynamic and go back to Boltzmann and Einstein. Binomiality and toricity have been widely studied in mathematics [13, 29, 10]. Binomiality is a hard problem and the typical approach by computing a Gröbner basis is \text{EXPSPACE}-complete [21].

In CRN theory, various articles have been written on binomiality and toricity, e.g., by Gorban et al. [14] and Grigoriev and Weber [17] and also other authors [9, 27]. Feinberg [11] and Craciun, et al. [7] have studied toric dynamical systems.

Dickenstein et al. have presented sufficient linear algebra conditions with inequalities for conditional binomiality in [24], which lead to MESSI CRNs [23]. For homogeneous ideals, it has been shown that Dickenstein et al.’s condition is necessary as well [6]. The latter has been implemented in Maple and Macaulay II in [19, 18].

A geometric view towards toricity has been considered by Grigoriev et al. in [16], introducing shifted toricity and presenting algorithms using quantifier elimination [8, 15, 30] and Gröbner bases [3, 4]. A first order logic test for toricity has been given in [25]. In [26] unconditional binomiality has been introduced and for reversible reactions a polynomial time algorithm, based on a linear algebra approach, has been presented.

Graph theory has been used in the study of CRNs, e.g., for detecting **concordance** and **weak monotonicity** [12, 28].

The main contribution of this article is a graph theoretical approach for testing unconditional binomiality of an RCRN. We use a modification of **species–reaction** graphs and present an algorithm equivalent to the linear algebra approach presented in [26]. We have implemented the graph theoretical algorithm as well the linear algebra algorithm in Maple [20] and compared them with each other and with the algorithms based on Gröbner basis and quantifier elimination presented in [16] via experiments on biological models from the BioModels repository [5]. Our experiments showed that the graph theoretical and linear algebra approaches are not only drastically faster, but they can also handle many cases that were not possible to compute using Gröbner basis and quantifier elimination.

The plan of this article is as follows. In Section 2 we briefly review the linear algebra approach in [26]. Section 3 presents the graph theoretical algorithm and the proof of its correctness, which are the main contributions of this article. Implementations of and experiments using both the graph theoretical and linear algebra algorithms are presented in Section 4.

## 2 Testing Unconditional Binomiality via Linear Algebra

In this section we briefly review the linear algebra method in [26] for testing unconditional binomiality in RCRNs.

**Definition 1** (Definition 1, [26]). Let \( C \) be a reversible chemical reaction network with the complexes \( C_1, \ldots, C_s \), let \( k_{ij}, 1 \leq i \neq j \leq s \), be the constant rate of the reaction from \( C_i \) to \( C_j \), and let \( x_1, \ldots, x_n \) be the concentrations of the species. \( b_{ij} := -k_{ij}m_i + k_{ji}m_j \) is called the binomial associated to the reaction from \( C_i \) to \( C_j \).

If there is no reaction between \( C_i \) and \( C_j \), set \( b_{ij} := 0 \).

**Example 1.** The binomials associated to the reversible reactions presented in the introduction section are \( b_{12} = -k_{12}x_1x_2 + k_{21}x_3 \) and \( b_{23} = -k_{23}x_3 + k_{32}x_1x_4 \).

Following Definition 1, one can write the corresponding ODEs for an RCRN as

\[
\dot{x}_k = p_k = \sum_{j=1}^{s} a_{ij}^{(k)} b_{ij}, \quad k = 1, \ldots, n, \quad (2)
\]
where \( c_{ij}^{(k)} \) is the difference between the stoichiometric coefficients of the \( k \)-th species in the reaction \( C_i \rightleftharpoons C_j \). It has been shown that if rate constants are indeterminates then the above sum of binomials representation is unique [26].

**Definition 2** (Definition 6, [26]). Let \( C \) be an RCRN as in Definition 1 and assume that \( p_i \), the generators of its steady state ideal, are written as sum of binomials as in Equation 2. We define the binomial coefficient matrix of \( C \) to be the matrix whose rows are labeled by \( p_1, \ldots, p_n \) and whose columns are labeled by nonzero \( b_{ij} \) and the entry in row \( p_k \) and column \( b_{ij} \) is \( c_{ij}^{(k)} \).

The binomial coefficient matrix can be used to test the unconditional binomiality of an RCRN.

**Theorem 3** (Theorem 8, [26]). A RCRN is unconditionally binomial (i.e., assuming the rate constants to be indeterminates) if and only if the reduced row echelon form (RREF) of its binomial coefficient matrix has at most one nonzero entry at each row.

The theorem leads to a linear algebra approach for testing unconditional binomiality [26, Algorithm 1], which is polynomial time in terms of the size of matrix, which is the maximum of the number of species and reactions in the RCRN. In Section 3, we present an implementation of the algorithm in Maple and compare it with the other approaches.

### 3 A Graph Theoretical Approach

In this section we present a graph theoretical algorithm equivalent to the linear algebra approach in [26]. A CRN can be represented as a graph in several ways. A first idea is the well-known complex–reaction graph of a CRN, in which the complexes are considered as the vertices and the reactions as the directed edges. Another idea is species–reaction graphs, which is used to study concordance and weakly monotonicity of kinetics in a CRN [12].

![Figure 1: Mod. Spec–Reac. Graph (Example 2)](image)

**Definition 4** (Modified Species–Reaction Graph of an RCRN). Let \( S \) be the set of species and \( R \) the set of reactions of a given RCRN. The modified species–reaction graph \( G \) of the RCRN is defined as follows.

- For each species \( s \in S \), consider a vertex of \( G \) (species vertices)
- For each reaction \( r \in R \), consider a vertex of \( G \) (reactions vertices)
- For each reaction vertex, there exists an undirected edge to the species vertices of the species that appear in the reaction.
- Each edge of the graph is labeled by an integer number which is the difference between the stoichiometric coefficients of the species (present at one end of the edge) in the reactant and product complexes.

**Example 2.** The species–reaction graph of the following RCRN is showed in Figure 1:

\[
2A + B \rightleftharpoons C \rightleftharpoons A \rightleftharpoons 2B.
\]

In order to check unconditional binomiality of an RCRN using the modified species–reaction...
graph, we present an algorithm that modifies the graph by adding and deleting the edges, and updating the edge labels so that unconditional binomiality can be easily read off from the final graph. The algorithm simulates the procedure described in Theorem 3 from Section 2 and Algorithm 1 from 26 by reducing a binomial coefficient matrix to its RREF in order to test unconditional binomiality of the network.

The idea of the algorithm is as follows. First, we iterate through the reaction vertices, selecting and marking an (random) unmarked species vertex connected to the current reaction vertex (initially all species vertices are unmarked). If there are no unmarked species vertices, then the current reaction vertex is skipped. Then, we delete all the edges incident to the current reaction vertex that are different from the edge going to the marked species vertex. Furthermore, if an edge exists from the marked species vertex to a reaction vertex, then we add the edge from the reaction vertex to the current species vertex. Nevertheless, if the edges already exists, then we update the label accordingly and if the new label is zero then we eliminate the edge. The final graph is reached when all reaction vertices have been visited. At the end of the algorithm, unconditional binomiality is checked by testing if each component of the final graph contains either only one species vertex or one species vertex and one reaction vertex. The algorithm is fully described below.

The functions in Algorithm 1 are self-explanatory. We just mention that ElimEdge eliminates the edge connecting a species vertex and reaction vertex, and UpdCf updates the coefficient of the edge that goes from a species vertex to a reaction vertex (Detailed description of the functions can be found in the Git repository 22).

Theorem 5. Algorithm 1 is correct, terminates and its asymptotic worst case time complexity can be bounded by $O(\max(r,n)^2)$, where $\omega$ is the constant appearing in the complexity of matrix multiplication, $r$ is the number of reactions and $n$ is the number of species of an RCRN.
Proof. (Proof of the correctness) Assume that the algorithm output is unconditionally binomial, then we must prove that the steady state ideal of the RCRN is binomial. To do so, we will show that the steps of the graph theoretical Algorithm 1 are equivalent to the steps of the linear algebra approach in [26, Algorithm 1] using Reduced Row Echelon Form (RREF) in the binomial coefficient matrix. Based on Theorem 3, Algorithm 1 from [26] initially selects a pivot in the matrix which is equivalent to marking an unmarked species vertex that is connected to a reaction vertex in steps 9 to 14. Reducing the nonzero entries to zero in the column of the selected pivot in the matrix is equivalent to eliminating the edges from the reaction vertex to all other species vertices that are not the one selected in step 20. While performing the reduction, some entries of other rows may be affected. The equivalent to this in Algorithm 1 is the update of the coefficients in some edges in step 27 or the elimination of edges in step 29 or the addition of edges in step 32. Then, obtaining the RREF of the matrix is equivalent to a combination of the following items:

- species vertices without edges, which are equivalent to the zero rows and/or
- reaction(species) vertices connected to at least one species(reaction) vertex, which are equivalent to the final columns(rows) of the matrix.

Finally, testing unconditional binomiality in the matrix (via checking that it has at most one nonzero entry at each row) is equivalent to checking that the components of the final graph contain either:

- only one species vertex or,
- one species vertex and one reaction vertex connected to one another.

(Proof of termination) For the proof of termination, note that the graph has finitely many edges and vertices. Hence, each of the loops terminates at some point. The loops at lines 6 and 22 eventually terminate as they iterate through reaction vertices which are finite and there is no creation of such vertices anywhere. Likewise, the loops at lines 4, 10 and 17 terminate as they iterate through species vertices which are also finite, and no new vertices are created.

(Proof of the complexity bound) This comes from the fact that each operation in the graph theoretical algorithm corresponds to an operation in the linear algebra approach. □

As a sidenote, for any transformed graph one can generate a binomial coefficient matrix by taking the species vertices as rows, reactions vertices as columns and the labels of edges as entries and vice versa, from any binomial coefficient matrix one can generate a graph by applying the reverse procedure.

Example 3. Graphs generated at each step of Algorithm 1 for the following RCRN has been shown in Figure 2.

\[ 3B \rightarrow 2C + A \rightarrow 2D + 2B \rightarrow 3B. \]

The final graph has a component that contains two species vertices and three reaction vertices and therefore the RCRN is not unconditionally binomial. On the other hand, Algorithm 1 in [26] gives the following matrix which shows that the RCRN is not unconditionally binomial since the first and second rows contain more than one nonzero entries.

4 Experiments

We have implemented our graph theoretical approach in Algorithm 1 in Section 3 as well as the linear algebra approach in [26, Algorithm 1] in Maple. Both algorithms are available in the Git repository [22]. The performance of the implementations is tested on the biomodels from the BioModels repository [5] and the results are compared to the experiments done on the same biomodels using Gröbner basis and quantifier elimination in [16]. Note that in order to be able to test unconditional binomiality of biomodels,
we have assumed reversibility with free reverse rate constants, while in \[16\] binomiality with pre-assigned values of rate constants are tested. The result of the experiments can be found in the Git repository \[22\].

One can notice that overall, Algorithm \[1\] performed equal or better than \[26, Algorithm 1\]. For a vast amount of models the difference of the performance of those two algorithms is almost zero, which was predictable as the proof of Theorem \[5\] suggests that the steps in the graph theoretical algorithm and the linear algebra approach are equivalent. However, for some models (e.g., 205, 293 and 574) the graph theoretical approach is faster. It is worth noting that these models have large binomial coefficient matrices, which may suggest that for RCRN with a large number of species and/or reactions the graph theoretical approach can be faster. This will be investigated further in future as the number of models with large matrices is not enough for a thorough comparison at this stage.

Comparing Algorithm \[1\] with the Gröbner basis and quantifier elimination computations in \[16\] shows that for the great majority of the cases, the graph theoretical approach is much faster. More importantly, there are twenty biomodels that Gröbner basis and/or quantifier elimination methods in \[16\] run out of time (for a six hour limit of computations), while those cases were handled in less than three seconds via both graph theory and linear algebra approaches. For six biomodels Gröbner basis computations terminate in less than six hours, however, our graph theory algorithm as well as the linear algebra approach are at least 1000 times faster than those. For ten biomodels the graph theory (and the linear algebra) approach is at least 500 times faster than the computations via Gröbner basis and quantifier elimination.

An interesting observation is the quite high number (sixty nine) of the biomodels that are not considered in \[16\] for the unclear numeric values of their rate constants, whereas our graph theoretical (and linear algebra) approaches on almost all of those cases terminated in less than a second.

A comparison of the performance of our graph theory algorithm with the algorithms proposed in \[24, 6\] are similar to the performance of the linear algebra algorithm in \[26\] vs algorithms in \[24, 6\]. This is because, as it is mentioned earlier in this section, the graph theoretical algorithm has a similar (or better) performance to the linear algebra algorithm in \[26\]. In particular, for two reversible biomodels in the database (Biomodels 491 and 492), the graph theoretical method is more than twice faster than the linear algebra in \[26\], which means that it is much faster than the algorithms in \[24, 6\]. More details of some comparisons between the linear algebra methods in the literature for testing unconditional binomiality and conditional binomiality can be found in \[26, Section 3\].

5 Conclusion

The present work introduces an efficient graph theoretical algorithm for testing unconditional
binomiality in an RCRN, which is different from the conditional binomiality considered by several other authors. The algorithm is essentially equivalent to the linear algebra approach presented earlier for testing unconditional binomiality. Implementations of the graph theoretical algorithm as well the linear algebra approach are done in Maple and experiments are carried on over several biomodels.

While the graph theoretical algorithm seem to have a slight advantage over the linear algebra approach, the experiments reveal drastic advantage of both of those methods over the existing algorithms based on Gröbner basis and quantifier elimination. Additionally, many cases that could not be handle by the Gröbner basis and quantifier elimination methods in a reasonable time, were tested in less than few seconds via the graph theoretical approach.

Acknowledgments
This work has been supported by the bilateral project ANR-17-CE40-0036/DFG-391322026 SYMBIONT [1, 2].

References
[1] F. Boulier, F. Fages, O. Radulescu, S. Samal, A. Schuppert, W. Seiler, T. Sturm, S. Walcher, and A. Weber. The SYMBIONT project: Symbolic methods for biological networks. ACM Communications in Computer Algebra, 52(3):67–70, December 2018.
[2] F. Boulier, F. Fages, O. Radulescu, S. Samal, A. Schuppert, W. Seiler, T. Sturm, S. Walcher, and A. Weber. The SYMBIONT project: Symbolic methods for biological networks. F1000Research, 7(1341), 2018.
[3] Bruno Buchberger. Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal. Doctoral dissertation, Mathematical Institute, University of Innsbruck, Innsbruck, Austria, 1965.
[4] Bruno Buchberger. Ein algorithmisches kriterium für die lösbarkeit eines algebraischen gleichungssystems. Aequationes Mathematicae, 4:374–383, 01 1970.
[5] Vijayalakshmi Cheilliah, Nick Juty, Ishan Ajmera, Raza Ali, Marine Dumousseau, Mihai Glont, Michael Hucka, Gaël Jalowicki, Sarah Keating, Vincent Knight-Schrijver, Audal Lloret-Villas, Kedar Nath Nataraajan, Jean-Baptiste Pettit, Nicolas Rodríguez, Michael Schubert, Sarala M. Wimalaratne, Yangyang Zhao, Henning Hermjakob, Nicolas Le Novère, and Camille Laibe. BioModels: Ten-year anniversary. Nucl. Acids Res., 2015.
[6] Carsten Conradi and Thomas Kahle. Detecting binomiality. Advances in Applied Mathematics, 71:52–67, 2015.
[7] Gheorghe Craciun, Alicia Dickenstein, Anne Shiu, and Bernd Sturmfels. Toric dynamical systems. J. Symb. Comput., 44(11):1551–1565, November 2009.
[8] James H. Davenport and Joos Heintz. Real quantifier elimination is doubly exponential. J. Symb. Comput., 5(1–2):29–35, February–April 1988.
[9] Alicia Dickenstein, Mercedes Pérez Millán, Anne Shiu, and Xiaoxian Tang. Multistationarity in structured reaction networks. Bulletin of Mathematical Biology, 81(5):1527–1581, Feb 2019.
[10] David Eisenbud and Bernd Sturmfels. Binomial ideals. Duke Mathematical Journal, 84(1):1–45, Jul 1996.
[11] Martin Feinberg. Complex balancing in general kinetic systems. Arch. Ration. Mech. An., 49(3):187–194, January 1972.
[12] Martin Feinberg. Foundations of Chemical Reaction Network Theory, volume 202
of *Applied Mathematical Sciences.* Springer, 2019.

[13] William Fulton. *Introduction to Toric Varieties,* volume 131 of *Annals of Mathematics Studies.* Princeton University Press, 1993.

[14] Alexander N. Gorban and Vassili N. Kolokoltsov. Generalized mass action law and thermodynamics of nonlinear Markov processes. *Math. Model. Nat. Phenom.*, 10(5):16–46, 2015.

[15] Dima Grigoriev. Complexity of deciding Tarski algebra. *J. Symb. Comput.*, 5(1–2):65–108, February–April 1988.

[16] Dima Grigoriev, Alexandru Iosif, Hamid Rahkooy, Thomas Sturm, and Andreas Weber. Efficiently and effectively recognizing toricity of steady state varieties. *CoRR*, abs/1910.04100, 2019.

[17] Dima Grigoriev and Andreas Weber. Complexity of solving systems with few independent monomials and applications to mass-action kinetics. In *Proceedings of the CASC 2012*, volume 7442 of *LNCS*, pages 143–154. Springer, 2012.

[18] Alexandru Iosif and Hamid Rahkooy. Analysis of the Conradi-Kahle algorithm for detecting binomiality on biological models. *CoRR*, abs/1912.06896, 2019.

[19] Alexandru Iosif and Hamid Rahkooy. Maplebinomials, a Maple package for testing binomiality, August 2019. Available at https://doi.org/10.5281/zenodo.3564428.

[20] Maplesoft, a division of Waterloo Maple Inc., Waterloo, Ontario, 2019.

[21] Ernst W. Mayr and Albert R. Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Adv. Math.*, 46(3):305–329, 1982.

[22] Cristian Vargas Montero and Hamid Rahkooy. TestBinomialityRCRN, Maple modules for testing unconditional binomiality of reversible chemical reaction networks, June 2020. Available at https://doi.org/10.5281/zenodo.3901559.

[23] Mercedes Pérez Millán and Alicia Dickenstein. The structure of MESSI biological systems. *SIAM J. Appl. Dyn. Syst.*, 17(2):1650–1682, 2018.

[24] Mercedes Pérez Millán, Alicia Dickenstein, Anne Shiu, and Carsten Conradi. Chemical reaction systems with toric steady states. *Bull. Math. Biol.*, 74(5):1027–1065, October 2012.

[25] Hamid Rahkooy and Thomas Sturm. First-order tests for toricity. *CoRR*, abs/2002.03586, 2020.

[26] Hamid Rahkooy and Thomas Sturm. A linear algebra approach for detecting binomiality of steady state ideals of reversible chemical reaction networks. *CoRR*, abs/2002.12693, 2020.

[27] AmirHosein Sadeghimanesh and Elisenda Feliu. The multistationarity structure of networks with intermediates and a binomial core network. *Bulletin of Mathematical Biology*, 81(7):2428–2462, May 2019.

[28] Guy Shinar and Martin Feinberg. Concordant chemical reaction networks and the species-reaction graph. *Mathematical biosciences*, 241, 08 2012.

[29] Bernd Sturmfels. *Gröbner bases and convex polytopes,* volume 8 of *University Lecture Series.* AMS, Providence, RI, 1996.

[30] Volker Weispfenning. The complexity of linear problems in fields. *J. Symb. Comput.*, 5(1–2):3–27, February–April 1988.

[31] Douglas B. West. *Introduction to Graph Theory.* Pearson Modern Classics for Advanced Mathematics Series. Pearson, 2017.