THE GAPS IN ROGERS' PROOF
OF IDENTIFIABILITY OF THE
GTR+Γ+I MODEL

ELIZABETH S. ALLMAN, CÉCILE ANÉ, AND JOHN A. RHODES

This note explains the gaps in the published proof of Rogers [Rog01] that the GTR+Γ+I model is identifiable. Since that paper has been widely cited and accepted as correct, our goal is to clearly indicate where the argument is flawed, and illustrate, through some examples, the nature of the logical gaps.

We emphasize that we do not *prove* that the gaps in the published argument cannot be bridged. Indeed, it seems most likely that the GTR+Γ+I model is identifiable, at least for generic parameters, and it is possible a correct proof might follow the rough outline of [Rog01]. However, we have not been able to complete the argument Rogers attempts to make. Moreover, our own proof of the identifiability of the GTR+Γ model (manuscript in preparation) follows a different line of argument.

1. Gaps in the published proof

There are two gaps in Rogers' argument which we have identified. In this section we indicate the locations and nature of these flaws, and in subsequent ones we elaborate on them individually.

The first gap in the argument occurs roughly at the break from page 717 to page 718 of the article. To explain the gap, we first outline Rogers' work leading up to it. Before this point, properties of the graph of the function $\nu^{-1}(\mu(x))$ have been carefully derived. An example of such a graph, for particular values of the parameters α, a, π, p occurring in the definitions of ν and μ, is shown in Figure 2 of the paper. For these parameter values and others, the article has carefully and correctly shown that for $x \geq 0$ the graph of $\nu^{-1}(\mu(x))$

(1) is increasing,
(2) has a single inflection point, where the graph changes from convex to concave (i.e, the concavity changes from upward to downward),
(3) has a horizontal asymptote as $x \to \infty$.

Although the article outlines other cases for different ranges of the parameter values, Rogers highlights the case when these three properties hold.

At the top of page 718 of the article, Figure 3 is presented, plotting the points whose coordinates are given by the pairs $(\nu^{-1}(\mu(\tau_1 \lambda_i)), \nu^{-1}(\mu(\tau_2 \lambda_i)))$ for all $\lambda_i \geq 0$. Here $\tau_2 > \tau_1$ are particular values, while α, a, π, p are given the values leading to Figure 2. Rogers points out that "As in Figure 2, the graph [of Figure 3] has an inflection point, is concave upwards before the inflection point, and is concave downwards after the inflection point." Then he claims that "Similar graphs will be produced for any pair of path distances such that $\tau_2 > \tau_1$." However, he gives no argument for this claim. As the remainder of the argument strongly uses the

concavity properties of the graph of his Figure 3 (in the second column on page 718 the phrase "...as shown by Figure 3" appears), without a proof of this claim the main result of the paper is left unproved.

Judging from the context in which it is placed, a more complete statement of the unproved claim would be that for any values of α, a, π, p resulting in a graph of $\nu^{-1}(\mu(x))$ with the geometric properties of Figure 2, and any $\tau_2 > \tau_1$, the graph analogous to Figure 3 has a single inflection point. As no argument is given to establish the claim, we can only guess what the author intended for its justification. From what appears earlier in the paper, it seems likely that the author believed the three geometric properties of the graph in Figure 2 enumerated above implied the claimed properties of Figure 3. However, that is definitely not the case, as we will show in Section 2 below.

Note that we do not assert that the graphs analogous to Figure 3 for various parameter values are not as described in [Rog01]. While plots of them for many choices of parameter values certainly suggest that Rogers' claim holds, it is of course invalid to claim a proof from examples. Moreover, with 4 parameters α, a, π, p to vary, it is not clear how confident one should be of even having explored the parameter space well enough to make a solid conjecture. In light of the example we give in Section 2, justifying Rogers' claim would require a much more detailed analysis of the functions ν and μ than Rogers attempts.

If this first gap in the proof were filled, a second problem would remain. Though less fundamental to the overall argument, this gap would mean that identifiability of the model would be established for *generic* parameters, but that there might be exceptional choices of parameters for which identifiability failed. ('Generic' here can be taken to mean for all parameters except those lying in a set of Lebesgue measure zero in parameter space. More informally, for any reasonable probability distribution placed on the parameter space, randomly-chosen parameters will be generic.)

Although the origin of this problem with non-generic parameters is clearly pointed out by Rogers, it is open to interpretation whether he attempts to extend the proof to all parameter values at the very end of the article. However, as the abstract and introductory material of [Rog01] make no mention of the issue, this point at the very least seems to have escaped many readers attention.

This gap occurs because the published argument requires that the non-zero eigenvalues of the GTR rate matrix Q be three *distinct* numbers. On page 718, at the conclusion of the main argument, it is stated that "Therefore, if the substitution rate matrix has three distinct eigenvalues, the parameters of the I+Γ rate heterogeneity will be uniquely determined." The author then goes on to point out that for the Jukes-Cantor and Kimura 2-parameter models this assumption on eigenvalues is violated, but "[f]or real data sets, however, it is unlikely that any two or all three of the eigenvalues will be exactly identical."

Leaving aside the question of what parameters one might have for a model which fits a real data set well, Rogers here clearly indicates that his proof of identifiability up to this point omits some exceptional cases. In the concluding lines of the paper, he points out that these exceptional cases can be approximated arbitrarily closely by parameters with three distinct eigenvalues. While this is true, such an observation cannot be used to argue that the exceptional cases are not exceptional, as we will

discuss below in Section 3. It is unclear whether the concluding lines of [Rog01] were meant to 'fill the gap' or not.

Of course, one might not be too concerned about exceptional cases. Indeed, if the first flaw were not present in his argument, then Rogers' proof would still be a valuable contribution in showing that for 'most' parameter values identifiability held. One might then look for other arguments to show identifiability also held in the exceptional cases. Nonetheless, it is disappointing that the exceptional cases include models such as the Jukes-Cantor and Kimura 2-parameter that are well-known to biologists and might be considered at least reasonable approximations of reality in some circumstances.

2. A COUNTEREXAMPLE TO THE GRAPHICAL ARGUMENT

It seems that the origin of the first flaw in Rogers' argument is in a belief that the three enumerated properties he proves are exhibited in his Figure 2 result in the claimed properties of his Figure 3. In this section, we show this implication is not valid, by exhibiting a function whose graph has the three properties, but when the graph analogous to Figure 3 is constructed, it has multiple inflection points.

Let
$$f(x) = \int_0^x \exp\left(\exp\left(-10(t-1)^2\right) - \frac{(1-t)^2}{10}\right) dt.$$
Then $f(0) = 0$, and
$$f'(x) = \exp\left(\exp\left(-10(x-1)^2\right) - \frac{(1-x)^2}{10}\right),$$
so $f'(x) > 0$ and f is increasing. Furthermore, one sees that $f'(x)$ decays quickly enough to 0 as $x \to \infty$, so that $f(x)$ has a horizontal asymptote as $x \to \infty$.

To see that $f(x)$ has a single inflection point where the graph passes from convex to concave, it is enough to show $f'(x)$ has a unique local maximum and no local minima. But this would follow from $g(x) = \ln(f'(x))$ having a unique local maximum and no local minima. Since
$$g(x) = \exp\left(-10(x-1)^2\right) - \frac{(1-x)^2}{10},$$
and the two summands here have unique local maxima at $x = 1$ and no local minima, g must as well. Thus f exhibits the enumerated properties of Rogers' Figure 2. For comparison, we graph f in our Figure 1 below.
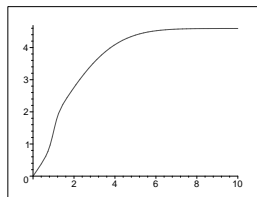

FIGURE 1. The graph $y = f(x)$.

The analog of Figure 3 for the function f would show the points $(f(\tau_1 x), f(\tau_2 x))$. If we choose $\tau_1 = 1, \tau_2 = 2$, we obtain the graph shown in our Figure 2. Obviously,
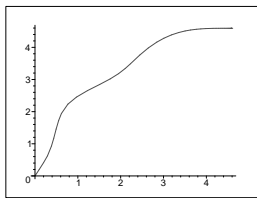
FIGURE 2. The points $(f(x), f(2x))$.

the curve in Figure 2 has multiple — at least three — inflection points. Although we will not give a formal proof here that this curve has multiple inflection points, it is not difficult to do so.

3. IDENTIFIABILITY FOR GENERIC PARAMETERS VS. ALL PARAMETERS

The second gap in Rogers' argument arises because it is possible to have identifiability for generic parameters, but not for all parameters. Even if identifiability of generic parameters has been proved, then one cannot easily argue that identifiability must hold for the non-generic, exceptional cases as well. To illustrate this, we give a simple example.

Consider the map $\phi : \mathbb{R}^2 \to \mathbb{R}^2$, defined by

$$\phi(a, b) = (a, ab).$$

Here a, b play the roles of 'parameters' for a hypothetical model, whose 'joint distribution' is given by the vector-valued function ϕ.

Suppose (x, y) is a particular distribution which arises from the model (i.e., is in the image of ϕ), and we wish to find a, b such that $\phi(a, b) = (x, y)$. Then provided $x \neq 0$ (or equivalently $a \neq 0$), it is straightforward to see that a, b must be given by the formulas

$$a = x, \quad b = y/x.$$

Thus for generic a, b (more specifically, for all (a, b) with $a \neq 0$) this hypothetical model is identifiable.

Notice, however, that if $(x, y) = (0, 0)$, the situation is quite different. From $x = 0$, we see that we must have $a = 0$. But since $\phi(0, b) = (0, 0)$, we find that all parameters of the form $(0, b)$ lead to the same distribution $(0, 0)$. Thus these exceptional parameters are *not* identifiable. Therefore, we have identifiability precisely for all parameters in the 2-dimensional ab-plane *except* those lying on the 1-dimensional line where $a = 0$. These exceptional parameters, forming a set of lower dimension than the full space, have Lebesgue measure zero within it.

Notice that even though there are parameter values arbitrarily close to the exceptional ones $(0, b)$ which are identifiable (for instance, (ϵ, b) for any small $\epsilon \neq 0$), it is invalid to argue that the parameters $(0, b)$ must be identifiable as well.

This example shows that even if the the first flaw in the argument of [Rog01] were repaired, the approach outlined there will at best give identifiability for generic parameters. The final lines of that paper are not sufficient to prove identifiability for all parameter values.

Obviously the function ϕ given here could not really be a joint distribution for a statistical model, since the entries of the vector $\phi(a, b)$ do not add to one, nor are they necessarily non-negative. However, these features can be easily worked into a more complicated example. If one prefers a less contrived example, then instances of generic identifiability of parameters but not full identifiability occur in standard statistical models used outside of phylogenetics (for instance, in latent class models). We have chosen to give this simpler example to highlight the essential problem most clearly.

References

[Rog01] James S. Rogers, *Maximum likelihood estimation of phylogenetic trees is consistent when substitution rates vary according to the invariable sites plus gamma distribution*, Syst. Biol. **50** (2001), no. 5, 713–722.

Department of Mathematics and Statistics, University of Alaska Fairbanks
E-mail address: e.allman@uaf.edu

Departments of Statistics and Botany, University of Wisconsin Madison
E-mail address: ane@stat.wisc.edu

Department of Mathematics and Statistics, University of Alaska Fairbanks
E-mail address: j.rhodes@uaf.edu