This Supplemental Material is divided into five sections, which provide (i) details of calculations for \( N = 2 \), (ii) the derivation of the semianalytic approach for \( N = 3 \) and a discussion of its numerical implementation, (iii) an analysis of the statistics of the wavenumber at which a Turing instability first arises, (iv) a discussion of Turing instabilities with nondiffusing “slow” species, and (v) a proof of the asymptotic result claimed in the conclusion of our Letter.

## I. DETAILS OF CALCULATIONS FOR \( N = 2 \)

### A. Derivation of Eq. (3)

The form of the condition for Turing instability in Eq. (3) follows from that in Eq. (2.26) on page 85 of Vol. II of Ref. [S1] which, in our notation, reads

\[
f_u + dg_v \geq 2\sqrt{dJ}, \quad (S1)
\]

a quadratic in \( d = d_u/d_v \). Hence

\[
\sqrt{d} \geq \sqrt{d_e} = \frac{\sqrt{J} + \sqrt{J - fg_u}}{g_v} \quad \text{if } g_v \geq 0. \quad (S2)
\]

We notice that Eq. (S1) requires \( f_u + dg_v \geq 0 \). Since \( I_1 < 0 \), this implies that \( d \geq 1 \) if \( g_v \geq 0 \). Hence \( D_2^* = d_1 \) if \( g_v > 0 \), but \( D_2^* = 1/d_1 \) if \( g_v < 0 \). Now, if \( g_v \geq 0 \), then \( |f_u| \geq |g_v| \) because \( I_1 < 0 \) and \( p > 0 \). Equation (3) then follows, since

\[
\frac{g_v}{\sqrt{J} - \sqrt{J + \rho}} = \frac{\sqrt{J} + \sqrt{J + \rho}}{f_u}. \quad (S3)
\]

### B. Mass-conserving reaction-diffusion systems

Mass-conserving reaction-diffusion systems [S2] have \( g(u, v) = -f(u, v) \), and hence \( g_v = -f_u, \ g_v = -f_v \). The necessary conditions for Turing instability [S1] reduce to the pair of conditions \( f_u - f_v < 0 \) and \( f_u f_v > 0 \), i.e. \( 0 < f_u < f_v \) or \( f_v < f_u < 0 \). From Eqs. (3) and (4), it then follows that \( D_2^* = \max \{|f_u|/|f_v|, |f_u|/|f_v|\} = R \).

### C. Derivation of Eq. (5)

Let \( I = [-R, -1] \cup [1, R] \). Equation (3) shows that \( D_2^* \) is continuous on \( I^1 \), so attains its maximum value on that domain. Since \( p > 0 \) and \( J > 0 \), \( q \equiv -f_v g_u > 0 \), so that \( J + p = q \). Now \( D_2^* \) only depends on \( f_v, g_u \) through \( q \), and, by direct computation from Eq. (3),

\[
\frac{\partial D_2^*}{\partial q} = \frac{D_2^*}{\sqrt{J(J + p)}} > 0. \quad (S4)
\]

Hence \( D_2^* \) increases with \( q \), so \( (f_v, g_u) = (\pm (R, -R) \text{ at the maximum.}

Now assume that \( |f_u| \geq |g_v| \). Since \( I_1 < 0 \) and \( |f_u| \geq |g_v| \), it follows that \( f_u < 0 \) and \( g_v > 0 \). Then

\[
\frac{\partial D_2^*}{\partial f_u} = \frac{\sqrt{J_1 + \sqrt{J_1}}}{g_v \sqrt{J_1}} > 0, \quad \frac{\partial D_2^*}{\partial g_v} = -\frac{(\sqrt{J_1 + \sqrt{J_1}})^3}{g_v^3 \sqrt{J_1}} < 0. \quad (S5)
\]

and so \( (f_u, g_v) = (1, -1) \) at the maximum. If \( |f_u| \leq |g_v| \), we similarly find that \( (f_u, g_v) = (-1, 1) \) at the maximum. These parameters have range \( R \), and substituting these values into Eq. (3) yields Eq. (5).

### D. Calculation of \( P(D_2^* \leq D) \) for \( D \leq R \)

There are 48 ways of assigning values \( \pm 1 \) and \( \pm R \) to two of the entries \( f_u, f_v, g_u, g_v \) of \( J \). Integrating the conditions for Turing instability of the remaining entries in each of these cases using MATHEMATICA (Wolfram, Inc.) gives the area of parameter space in which a Turing instability arises,

\[
\int_{-1}^{1} \int_{R}^{J} \int_{0}^{p} \int_{1}^{\max |J|} \int_{1}^{\min |J|} \left( J > 0, I_1 < 0 \right) \left( p > 0 \right) \left( \max |J| = R \right) \left( \min |J| = 1 \right) \left( J - 1 \right)^2 \text{dJ} = 12(R - 1)^2, \quad (S6)
\]

where we use the shorthand \( \text{dJ} = d f_u \text{d} f_v \text{d} g_u \text{d} g_v \). To analyze the condition \( D_2^* < R \), we note that the expression for \( D_2^* \) in Eq. (3) shows that we may swap \( f_u, g_v, \) and \( f_v, g_u \). Hence the 48 cases reduce to 4 cases (corresponding to the entries \( \pm 1 \) or \( \pm R \) being on the the same or on different diagonals):

1. \( |f_u| = R, |g_v| = 1 \); 2. \( |f_v| = R, |g_u| = 1 \);
3. \( |f_u| = R, |f_v| = 1 \); 4. \( |f_u| = 1, |f_v| = R \).

Moreover, since \( q > 0 \), we may take \( f_v > 0 \) and \( g_u < 0 \) without loss of generality. We now discuss these cases separately.

1. \( I_1 < 0 \) implies \( f_u = -R, g_v = 1 \), and so

\[
D_2^* = \left( \sqrt{q} + \sqrt{q - R} \right)^2 \geq R. \quad (S7)
\]
(2) \( f_u g_v = -R \) since \( q > 0 \), so \( J = f_u g_v + R \).

(3) \( f_u = -R \) because \( I_1 < 0 \). Now \( p, q > 0 \), and so \( 0 < J = -R|g_v| - |g_u| < 0 \). This is a contradiction.

(4) \( f_u = 1 \) as \( I_1 < 0 \). Since \( g_v \leq 1 \), it follows that

\[
D_2^* = \left( \sqrt{-g_v R} + \sqrt{g_v R} - g_v \right)^2 \geq R. \tag{S8}
\]

In this way, \( D_2^* < R \) quantifies the diffusive threshold in a natural way. In particular, \( D_2^* < R \) is only possible in case (2). Since \( J > 0 \), we require \( f_u g_v + R > 0 \) in that case. Now \( I_1 < 0 \) and \( p > 0 \), so \( 1 < f_u < -R/g_v \) or \( 1 < g_v < -R/f_u \) depending on \( f_u > 0, g_v < 0 \) or \( f_u < 0, g_v > 0 \). Assume without loss of generality that \( |f_u| \geq |g_v| \). Then \( f_u < 0, g_v > 0 \) as \( I_1 < 0 \). Moreover, using Eq. (3), \( D_2^* = R \) if and only if \( g_v = 2 + f_u/R \). From Eqs. (S5), \( D_2^* \) decreases as \( g_v \) increases. Hence

\[
D_2^* < R \iff 2 + f_u/R < g_v \iff -R/f_u and f_u + g_v < 0, \tag{S9}
\]

using the conditions derived previously. Note that \(-R/f_u < R \) and \( 2 + f_u/R > 1 \) for \(-R < f_u < -1 \). If \( |f_u| < |g_v|, f_u, g_v \) are swapped in these conditions. Moreover, since \( q > 0 \), case (2) corresponds to 4 of the 48 cases. Hence we obtain, again using Mathematica,

\[
\int_1^{I_1} \int_{H_1} \int_{I_4} \left( \begin{array}{c}
J > 0, I_1 < 0 \\
p > 0, D_2^* < R \\
\max |J| = R \\
\min |J| = 1
\end{array} \right) \, dJ = 4 \left( \frac{2R(1-R)}{1+R} + R \log R \right). \tag{S10}
\]

Equations (S6) and (S10) imply

\[
\mathbb{P}(D_2^* < R) = \frac{R \left[ (R+1) \log R - 2(R-1) \right]}{3(R-1)^2(R+1)}. \tag{S11}
\]

In particular, \( \mathbb{P}(D_2^* < R) = O(\log R/R) \ll 1 \) for \( R \gg 1 \). This statement expresses the existence of the diffusive threshold mathematically.

From a more physical point of view, as discussed in our Letter, it is more natural to consider the probability \( \mathbb{P}(D_2^* < D) \), for some constant \( D > 1 \). Since “small” values \( R \ll D \) require fine-tuning of the reaction kinetics, we restrict to \( D \leq R \), so that \( D_2^* < D \) is only possible in case (2) above. We consider again the case \( g_v > 0, f_u < 0 \). Similarly to the derivation of conditions (S9), we find

\[
D_2^* < D \iff g_v > \sqrt{\frac{R}{D}}, \quad \frac{R}{g_v} \leq f_u < D g_v - 2\sqrt{D R},
\]

and \( f_u + g_v < 0 \). \tag{S12}

In particular,

\[
\frac{R}{g_v} = \max \left\{ -\frac{R}{g_v}, \frac{R}{g_v} \right\} \leq f_u \leq \min \left\{ -1, g_v, D g_v - 2\sqrt{D R} \right\}, \tag{S13a}
\]

in which, since \( g_v > 1 \),

\[
\min \left\{ -1, -g_v, D g_v - 2\sqrt{D R} \right\} = \begin{cases} -g_v & \text{if } g_v > 2\sqrt{D R} \over D + 1 \\
D g_v - 2\sqrt{D R} & \text{otherwise}, \end{cases} \tag{S13b}
\]

We notice that \( \sqrt{D} > 2\sqrt{D R}/(D+1) > \sqrt{R}/D \) since \( D > 1 \), and also that \( D g_v - 2\sqrt{D R} > -R/g_v \iff \sqrt{D g_v - 2\sqrt{D R}} > 0 \), but \(-g_v/R/g_v \iff g_v < \sqrt{R} \). The area of parameter space described by conditions (S12) is therefore

\[
\int_{2\sqrt{D R}}^{\sqrt{D}} \int_{f_u}^{g_v} \, df_u \, dg_v + \int_{\sqrt{D}}^{2\sqrt{D R}} \int_{f_u}^{-g_v} \, df_u \, dg_v
\]

\[
= \frac{R}{2} \log D - \frac{D-1}{D+1} R. \tag{S14a}
\]

Hence \( \mathbb{P}(D_2^* < D) = \frac{R}{3(R-1)^2} \left[ \log D - \frac{2(D-1)}{D+1} \right] \). \tag{S15}

E. Nondimensionalization

We close by remarking on the (absence of) nondimensionalization of the reaction system. Indeed, up to rescaling time, one among \( f_u, g_v, g_u, g_v \) can be set equal to \( \pm 1 \). Moreover, one more parameter can be set equal to \( \pm 1 \) by rescaling \( u, v \) differently. However, if we made those choices, we could no longer sample from a fixed interval.

II. SEMIANALYTIC METHOD FOR \( N = 3 \)

A. Derivation of the semianalytic method

1. Preliminary observations

Before deriving the semianalytic method, we need to make two preliminary observations.

First, the necessary and sufficient (Routh–Hurwitz) conditions for the homogeneous system to be stable include

\[
I_1 \equiv \tr J < 0 \quad \text{and} \quad J \equiv \det J < 0 \tag{S11}\]

By definition, \( \tilde{J}(k_z^2) \) has one zero eigenvalue. The other two eigenvalues are either real or two complex conjugates \( \lambda, \lambda^* \). In the second case, they are both stable (i.e. have negative real parts) since

\[
2 \Re(\lambda) = 0 + \lambda + \lambda^* = \tr \tilde{J}(k_z^2) = I_1 - k_z^2 \tr D < I_1 < 0. \tag{S16}
\]
Hence Eqs. (7) are not unstable to an oscillatory (Turing–Hopf) instability at \((d_u^*, d_v^*)\), so, by minimality of \((d_u^*, d_v^*)\), the system destabilizes to a Turing instability there.

Moreover, since \(\mathcal{J}\), viewed as a polynomial in \(k_2^2\), has leading coefficient \(-d_u \partial d_u\) and constant term \(\mathcal{J}(0) = J < 0\), the double root \(K(d_u, d_v)\) varies continuously with \(d_u, d_v\) and cannot change sign on a branch of \(\Delta(d_u, d_v) = 0\) in the positive \((d_u, d_v)\) quadrant.

2. Reduction of problem (9) to polynomial equations

The discriminant of \(\mathcal{J}\), viewed as a polynomial in the two variables \(d_u, d_v\), is

\[
\Delta(d_u, d_v) = \sum_{m=0}^{4} \sum_{n=0}^{4} \delta_{mn} d_u^m d_v^n, \quad (S17)
\]

where \(\delta_{00} = \delta_{10} = \delta_{01} = \delta_{34} = \delta_{43} = \delta_{44} = 0\) and (complicated) expressions for the 19 non-zero coefficients can be found in terms of the entries of \(J\) using Mathematica (Wolfram, Inc.).

The second remark above implies that, at a local minimum of \(D_3(d_u, d_v)\) on \(\Delta(d_u, d_v) = 0\), one of the following occurs:

(i) \(\Delta(d_u, d_v) = 0\) is tangent to a contour of \(D_3(d_u, d_v)\);

(ii) \(\Delta(d_u, d_v)\) intersects a vertex of a contour of \(D_3(d_u, d_v)\);

(iii) \(\Delta(d_u, d_v)\) is singular.

The contours of \(D_3(d_u, d_v)\) are drawn in Fig. 2(a) of our Letter and show that tangency to a contour in case (i) requires

\[
d d_u = 0 \quad \text{or} \quad d d_v = 0 \quad \text{or} \quad d d_v / d d_u = d_v / d_u. \quad (S18)
\]

Since \(\Delta(d_u, d_v) = 0\), the chain rule reads

\[
0 = d \Delta = \frac{\partial \Delta}{\partial d_u} d d_u + \frac{\partial \Delta}{\partial d_v} d d_v. \quad (S19)
\]

Hence there are two subcases:

(a) \(\frac{\partial \Delta}{\partial d_v} = 0\) or \(\frac{\partial \Delta}{\partial d_u} = 0\);

(b) \(d_u \frac{\partial \Delta}{\partial d_u} + d_v \frac{\partial \Delta}{\partial d_v} = 0\).

In subcase (a), \(\Delta\) viewed as a polynomial in \(d_v\) or \(d_u\) has a double root, and so its discriminant \([S4]\) must vanish. On removing zero roots, this discriminant of a discriminant is found to be a polynomial of degree 20 in \(d_u\) or \(d_v\), respectively; complicated expressions for its coefficients in terms of the non-zero coefficients \(\delta_{mn}\) in Eq. (S17) are obtained using Mathematica. Similarly, in subcase (b), the resultant \([S4]\) of \(\Delta\) and \(d_u \partial \Delta / \partial d_u + d_v \partial \Delta / \partial d_v\), viewed as polynomials in \(d_u\) or \(d_v\), must vanish. This resultant is another polynomial of degree 20 in \(d_v\) or \(d_u\).

Next, in case (ii), \(d_u = 1\) or \(d_v = 1\) or \(d_u = d_v\) [Fig. 2(a)], which reduces \(\Delta\) to three different polynomials in the single variable \(d_v, d_u\), or \(d = d_u = d_v\), respectively. These polynomials have degree 6.

Finally, in case (iii), we note that, at a singular point, \(\Delta = \partial \Delta / \partial d_u = \partial \Delta / \partial d_v = 0\), and so we are back in case (i), subcase (a).

Thus, we have reduced finding candidates for local minima in (9) to solving polynomial equations: this defines our semianalytic approach. The global minimum is found among those local minima with \(K(d_u, d_v) > 0\); in case (i), the roots only correspond to local minima if additionally \(d_u, d_v > 1\) or \(d_u, d_v < 1\) in subcase (a) and \(d_u < 1 < d_v\) or \(d_v < 1 < d_u\) in subcase (b) [Fig. 2(a)].

3. Extension to binary systems with \(N > 3\)

For binary systems, the diagonal entries of \(D\) take two different values, \(d_1, d_2\) only. Up to rescaling space, \(d_1 = 1\) and \(d_2 = d\), which turns the condition \(\Delta(D) = 0\) into \(2^{N-1} - 1\) different polynomial equations in the single variable \(d\), corresponding to the different combinatorial ways of assigning diffusivities \(d_1, d_2\) to the \(N\) species (in such a way that not all species have the same diffusivity). Determining the minimum value \(D_N^*\) of \(D_N\) = \(\max \{d, 1/d\}\) for these binary systems is thus reduced, again, to solving polynomial equations.

The argument we used above to show that coexistence of Turing and Turing–Hopf instabilities is not possible for \(N = 3\) does not, however, carry over to \(N > 3\). Numerically, it turns out, however, that systems in which Turing and Turing–Hopf instabilities coexist are rare. We therefore treat these systems in the same way as we treat systems for which the numerics fail (as discussed below).

B. Numerical implementation

Implementing the semi-analytical approach for \(N = 3\) and its extension to binary systems with \(4 \leq N \leq 6\) numerically takes some care as the coefficients of the polynomials that arise can range over many orders of magnitude. Our Python implementation therefore uses the mpmath library for variable precision arithmetic \([S5]\).

To determine the positive real roots of the polynomials that arise in the semi-analytical approach, we complement the Durand–Kerner complex root finding implemented in the mpmath library \([S5]\) with a test based on Sturm’s theorem \([S4]\), to ensure that all positive real roots are found. Those systems in which root finding fails—either because the Durand–Kerner algorithm fails to converge or because it finds an incorrect number of positive real roots—are discarded, but included in error estimates where reported.

C. Numerical samples

Table S1 gives the number of random Turing unstable systems from which distributions, averages, and probabilities were estimated for each \(R \in \{2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20\}\).

For \(N = 3\), we ran both a search for general, non-binary systems and a (larger but numerically less expensive) search for binary systems only. Since the first search only yielded
Asymmetric error bars again correspond to 95% confidence intervals larger than the plot markers, corrected for systems for which the numerics failed.

III. WAVENUMBER STATISTICS

In this Section, we discuss the wavenumber $k_N^*$, at which a Turing instability first arises at $D_N = D_N^*$. In particular, we discuss our Letter, we must ask whether a Turing instability is “observable at the system size”. This observability requires the lengthscale $1/k_N^*$ of the linear instability to be (a) smaller than the system size $L$ and (b) larger than $L/\ell$, for some scale difference $\ell > 1$. We are thus led to consider the probability $\mathbb{P}(K < k_N^* < \ell K)$, where $K = 1/L$.

It is instructive to start by considering the case $N = 2$. For the reaction-diffusion system in Eq. (1), a Turing instability arises for $D_2 = D_2^*$ at a wavenumber $k_2^* = (J/d_u d_v)^{1/4}$ [S1]. We stress that this value depends on $d_u, d_v$, not only through their ratio $d = d_u/d_v$. To absorb the dependence on the dimensional system scale, it is natural to consider

$$k_2(\ell) = \max_K \{ \mathbb{P}(K < k_2^* < \ell K) \},$$

(S20a)
as the maximal probability of a Turing instability being observable at some inverse system scale $K$ over a fixed scale difference $\ell$. We denote by $K_2(\ell)$ the corresponding maximizing inverse system size.

For $N > 2$, we correspondingly ask: what is the probability of a Turing instability being observable at this inverse system size? We therefore define

$$k_N(\ell) = \mathbb{P}(K_2(\ell) < k_N^* < \ell K_2(\ell))$$

(S20b)

for $N > 2$. We notice that $k_N(\ell)$ increases slightly with $N$. If we restrict the analysis to those Turing unstable systems with $D_N^* < D$, the probability is reduced somewhat for $N > 2$ compared to the case $N = 2$. This merely reflects the “fine-tuning problem”: the wavenumber is strongly constrained for those very rare systems that have a “small” diffusive threshold at $N = 2$. Moreover, the majority of the Turing instabilities at $N > 2$ do arise at physical wavenumbers, so we can extend the observations in Figs. 2(d) and 3(c) to note that random kinetic Jacobians are still more likely to be unstable to an observable Turing instability with small diffusive threshold for $N > 2$ than for $N = 2$.

IV. DIFFUSION OF “SLOW” SPECIES

In the notation of Eq. (12) of our Letter, Ref. [S6] shows that Turing instability at $d = 0$ requires $J_{22}$ to be stable (i.e. all its eigenvalues to have negative real part): if it is not, instabilities arise at arbitrarily small and therefore unphysical length scales. In particular, $\det J_{22} \neq 0$ and $J_{22}$ is invertible. Now, using another result of Ref. [S6],

$$\det (J - k^2D) = \det J_{22} \det (j - k^2I),$$

(S21)

where $j = J_{11} - J_{22}J_{12}J_{22}^{-1}$. Hence a Turing instability occurs at $d = 0$ only if $j$ has a positive real eigenvalue, as claimed in our Letter.

This also implies that a Turing instability at $d = 0$ requires $n > 2$ “fast” diffusers. Clearly, $n = 0$ is not possible. If $n = 1$, then Eq. (S21) yields $\det (J - k^2D) = \det J_{22}(j - k^2)$, where $j = J_{11} - J_{12}J_{22}^{-1}J_{21}$ is now a scalar. Using this result for $k = 0$ shows that $j \det J_{22} = \det J$. It follows that $\det (J - k^2D) = \det J - k^2 \det J_{22}$. Now $J$ is stable (for stability of the homogeneous steady state is a necessary condition for Turing instability), while $J_{22}$ is stable by the above. Hence sign $\det J = (-1)^n$ and sign $\det J_{22} = (-1)^{N-n} = (-1)^N$.

This shows that $\det J - k^2 \det J_{22} \neq 0$, and so $n \neq 1$.

V. THE ASYMPTOTIC DIFFUSIVE THRESHOLD

Let $J = O(1)$ be a Turing unstable kinetic Jacobian, with an eigenvalue $\lambda$ destabilizing at nearly equal diffusivities, so that $\Delta = 1 + d$ with $d = o(1)$. The following claim extends an argument of Ref. [S7]:

Claim. $J$ has a defective zero eigenspace.
Proof. Because $J - k^2I$ has a stable eigenvalue $\lambda - k^2$ and $-k^2d \ll J - k^2I$, the corresponding eigenvalue of

$$J - k^2D = (J - k^2I) - k^2d$$

can only have positive real part if $\lambda - k^2 = o(1)$ i.e. if $\lambda = o(1)$ and $k^2 = o(1)$ since Re($\lambda$) < 0. Hence J and J - k^2I have a zero eigenvalue at leading order. Additionally, the eigenvalue correction from $-k^2d = o(k^2)$ must be $O(k^2)$ at least, which occurs if and only if the (leading-order) zero eigenspaces of $J - k^2I$ and J are defective [S8]; this final implication is discussed in more detail in Ref. [S9].

The generic case is therefore $J = J_0 + O(\varepsilon)$, where $\varepsilon \ll 1$ and $J_0$ has a defective double zero eigenvalue.

Claim. $d \gtrsim O(\sqrt{\varepsilon})$; in particular, $D - I \gg J - J_0$.

Proof. Since $J_0$ has a defective double zero eigenvalue, J has two $O(\sqrt{\varepsilon})$ eigenvalues [S8], assumed to be stable (i.e. to have negative real parts). With $k = O(\varepsilon^*)$, $d = O(\varepsilon^d)$, destabilizing one of these requires, using the proof of the first claim above, $-k^2d \gtrsim O(\varepsilon)$ and $-k^2I \lesssim O(\sqrt{\varepsilon})$, i.e. $2\kappa + \delta \lesssim 1$ and $\kappa \gtrsim 1/2$. Hence $\delta \lesssim 1/2$. This proves the claim.

SUPPLEMENTAL CODE

The online Supplemental Material also includes excerpts from the Python code that we have written to implement the semianalytic approach for $N \geq 3$.

SUPPLEMENTAL REFERENCES

[S3] We were not able to compute this integral using Mathematica. This is why, by contrast with the derivation of Eq. (S10), we have given the full details of the calculation above.