Supplementary Information for the paper: On the effect of heterogeneity in stochastic interacting-particle systems

L. F. Lafuerza, R. Toral

Instituto de Física Interdisciplinar y Sistemas Complejos (IFISC), CSIC-UIB, Campus UIB, E-07122 Palma de Mallorca, Spain

A. M-states system

We consider here the case in which each particle can be in one of M (instead of 2) possible states. We will show that the results obtained in the main text for 2−state systems also hold in this more general case.

We label the states with the subscript $\alpha = 0, 1, \ldots, M-1$, so in this case the variable describing the state of particle i can take M possible values, $s_i = 0, \ldots, M-1$ (we start the labeling from 0 to be consistent with the previous case, that would correspond to $M = 2$). Let $p_i(\lambda_i, \alpha, t)$ the probability that particle i, with heterogeneity parameter λ_i, be on state α. It satisfies the evolution equation:

$$\frac{dp_i(\lambda_i, \alpha, t)}{dt} = \sum_{\beta} A_{\alpha,\beta}(\lambda_i) p_i(\lambda_i, \beta, t), \tag{1}$$

with $A_{\alpha,\beta}$ a general transition matrix (satisfying $\sum_{\gamma=0}^{N-1} A_{\gamma,\alpha} = 0$), that may depend in principle on time and on the time that the particle has been on its current state. To isolate the role of parameter heterogeneity, we assume that the initial condition is the same for all the particles (or that the initial condition is determined by the value of λ_i) such that the solution $p_i(\lambda_i, \alpha, t) = p(\lambda_i, \alpha, t)$ is the same for all particles sharing the same value of the parameter. The macroscopic state of the system will be described by the set of variables $n_\alpha = \sum_{i=1}^{N} \delta_{\alpha,s_i}$, that is, the number of particles in each state. The averages and variances of this variables are given by:

$$\langle n_\alpha(t) \rangle = \sum_{i=1}^{N} p(\lambda_i, \alpha, t) \tag{2}$$

$$\sigma^2[n_\alpha(t)] = \sum_{i=1}^{N} \left[p(\lambda_i, \alpha, t) - p(\lambda_i, \alpha, t)^2 \right]. \tag{3}$$

This variance is again smaller that tat of a system of identical particles with same average, the difference given by:

$$\sigma^2[n_\alpha(t)]_{\text{id}} - \sigma^2[n_\alpha(t)] = N\overline{p(\alpha,t)^2} - \overline{p(\alpha,t)}^2, \tag{4}$$

a result exactly analogous to the one obtained in the previous case. The heterogeneity among the particles on the probability of occupation of level α can be derived from the first moments of the occupation number of the level:

$$\overline{p(\alpha,t)^2} - \overline{p(\alpha,t)}^2 = \frac{\langle n_\alpha \rangle - \langle n_\alpha \rangle^2/N - \sigma^2[n_\alpha]}{N}. \tag{5}$$

Note that, when focusing on the number of particles on state α, the system effectively reduces to a 2−level one, with states α and no-α, so the results of the previous section can be translated directly.

A different and some times relevant question can be considered when the labeling of the states is such that the order is well defined (for example each state corresponds to an energy level or a distance from a reference). Then the average state is meaningful and we can study its statistical properties. Below we show that the variance of this mean level is again always smaller if heterogeneity is present.

The average state of the system is given by $L = \sum_{\alpha=0}^{M-1} \alpha \ \frac{n_\alpha}{N}$. It is a random variable whose average and variance are given by:

$$\langle L \rangle \ = \ = \sum_{\alpha=0}^{M-1} \alpha \frac{\langle n_\alpha \rangle}{N} = \sum_{\alpha=0}^{M-1} \sum_{i=1}^{N} \alpha \frac{p(\lambda_i, \alpha)}{N}, \tag{6}$$

$$\sigma^2[L] \ = \ \sum_{\alpha,\beta=0}^{M-1} \frac{\alpha\beta}{N^2} (\langle n_\alpha n_\beta \rangle - \langle n_\alpha \rangle \langle n_\beta \rangle) = \frac{1}{N^2} \sum_{i=1}^{N} \left[\sum_{\alpha=0}^{M-1} \alpha^2 p(\alpha, \lambda_i) - \sum_{\alpha,\beta=0}^{M-1} \alpha p(\alpha, \lambda_i) \beta p(\beta, \lambda_i) \right] \tag{7}$$

We have used $p(\lambda_i, \alpha) \ = \ \langle \delta_{\alpha, s_i} \rangle$ and $\langle n_\alpha n_\beta \rangle \ = \ \sum_{i,j=1}^{N} \langle \delta_{\alpha, s_i} \delta_{\beta, s_j} \rangle \ = \ \langle n_\alpha \rangle \langle n_\beta \rangle \ + \ \sum_{i=1}^{N} [\delta_{\alpha,\beta} p(\alpha, \lambda_i) - p(\alpha, \lambda_i) p(\beta, \lambda_i)]$. A system of identical particles that had the same average occupation of the different levels i.e. $p_{id}(\lambda_i, \alpha) = \frac{1}{N} \sum_{j=1}^{N} p(\lambda_j, \alpha,) = \frac{\langle n_\alpha \rangle}{N} \ \forall i, \alpha$, would have and average and variance of the mean level given by:

$$\langle L \rangle_{id} \ = \ \sum_{\alpha=0}^{M-1} \alpha \frac{\langle n_\alpha \rangle}{N} = \langle L \rangle, \tag{8}$$

$$\sigma^2[L]_{\rm id} \ = \ \frac{1}{N} \sum_{\alpha=0}^{M-1} \alpha^2 \frac{\langle n_\alpha \rangle}{N} - \frac{1}{N} \sum_{\alpha,\beta=0}^{M-1} \alpha\beta \frac{\langle n_\alpha \rangle}{N} \frac{\langle n_\beta \rangle}{N}. \tag{9}$$

We now define $g(\lambda_i) \equiv \sum_\alpha \alpha p(\lambda_i, \alpha)$ (the average level of particle i), and note that the first terms in the right-hand side of (7) and (9) are equal, while the second terms can be written as:

$$\frac{1}{N^2} \sum_{i=1}^{N} \sum_{\alpha,\beta=0}^{M-1} \alpha p(\lambda_i, \alpha) \beta p(\lambda_i, \beta) \ = \ \frac{1}{N^2} \sum_{i=1}^{N} g(\lambda_i)^2 = \frac{1}{N} \overline{g^2}, \tag{10}$$

$$\frac{1}{N} \sum_{\alpha,\beta=0}^{M-1} \alpha\beta \frac{\langle n_\alpha \rangle}{N} \frac{\langle n_\beta \rangle}{N} \ = \ \frac{1}{N} \left[\frac{1}{N} \sum_{i=1}^{N} g(\lambda_i) \right]^2 = \frac{1}{N} \overline{g}^2, \tag{11}$$

which implies that $\sigma^2[L]_{\rm id} \geq \sigma^2[L]$, i.e. the variance of the mean level is always smaller in a system of heterogeneous particles, the difference with respect to the case of identical ones

being:

$$\sigma^2[L]_{\rm id} - \sigma^2[L] = \frac{1}{N}\left(\overline{g^2} - \overline{g}^2\right) = \frac{1}{N}\sum_{\alpha,\beta=0}^{M-1}\alpha\beta\left[\sum_{i=1}^{N}\frac{p(\alpha,\lambda_i)p(\beta,\lambda_i)}{N} - \sum_{i,j=1}^{N}\frac{p(\alpha,\lambda_i)p(\beta,\lambda_j)}{N^2}\right] \geq 0. \tag{12}$$

The correction to the variance in this case scales as $1/N$, but again is of the same order as the variance itself, indicating a non-negligible correction. In this case to derive the heterogeneity of $g(\lambda_i)$ over the population one needs to know the average occupation level of each state $\langle n_\alpha \rangle$ and use:

$$\overline{g^2} - \overline{g}^2 = \sum_\alpha \alpha^2 \langle n_\alpha \rangle / N - \langle L \rangle^2 - N\sigma^2[L]. \tag{13}$$

This can be written in terms of the variance of L in an equivalent system of identical particles, $\sigma^2[L]_{\rm id}$. If this is known, one can directly use

$$\overline{g^2} - \overline{g}^2 = N\left(\sigma^2[L]_{\rm id} - \sigma^2[L]\right). \tag{14}$$

Note that, contrary to the two-level case, now the value of $\langle L \rangle$ does not determine $\sigma^2[L]_{\rm id}$.

B. Intuitive origin of the decrease of fluctuations for independent units

We have shown that a system of independent heterogeneous particles has smaller fluctuations for the collective variable than an equivalent system of identical ones. The origin of this result is the following (for simplicity we refer to the case of 2-state system):

The average of the global variable is determined by the concentration of the states of the particles around state 1 ($\langle n \rangle = \sum_i \langle s_i \rangle$). The fluctuations (measured by the variance) of the global variable are determined by the stochastic fluctuations of the individual particles alone ($\sigma^2[n] = \sum_i \sigma^2[s_i]$, since the particles are independent).

In a system of heterogeneous particles, the dispersion of the states of the particles is due to the heterogeneity (some prefer to be around sate 0, others prefer to be around sate 1) plus their intrinsic stochasticity. In a system of identical particles, the dispersion comes from the stochasticity alone, so for a system of identical particles to have the same concentration in the states of the particles (global average) than a heterogeneous system, the intrinsic stochasticity has to be larger. This will give rise to larger fluctuations for the global variable. In particular, any given rational value of $\frac{\langle n \rangle}{N} = \frac{A}{B}$ can be obtained with zero fluctuations, taking A particles that are always at state 1 and $B - A$ particles that are always at state 0.

This explanation is illustrated in figure (1). In the identical-particles system both particles fluctuate between 1 and 0. In the heterogeneous case, one particle spends most of the time at 1 and the other spends most of the time at 0. The probability of finding a given particle at 1 is the same in both cases (1/2) but in the heterogeneous case most of the time there is one particle at 1 and one particle at 0, resulting on a value of the average state most often equal to 1/2, and so with smaller fluctuations. The situation is similar for a larger

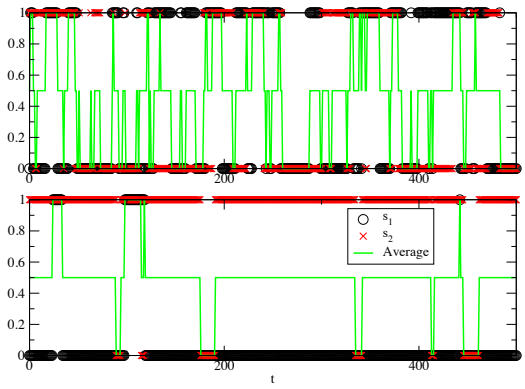

FIG. 1: Time series of a system of two identical (upper panel) and heterogeneous (lower panel) particles, together with the corresponding average state. Note that the fluctuations of the average state are more pronounced in the case of the identical particles.

number of particles, as shown in figure 1 of the main text. An analogous picture emerges when one considers more that 2 states. Note that in every case we compare a system of heterogeneous particles with another of identical ones that has the same one-particle distribution i.e. $p_i(\alpha)_{id} = \sum_j \frac{p(\alpha,\lambda_j)}{N}$, $\forall i, \alpha$.

C. Justification of the Ansatz

The general evolution equations for the first moments are of the form:

$$\frac{d\langle s_i\rangle}{dt} = \langle r_i^+\rangle - \langle (r_i^- + r_i^+)s_i\rangle, \tag{15}$$

$$\frac{d\langle s_i s_j\rangle}{dt} = -\langle q_{ij}s_j s_i\rangle + \langle r_i^+ s_j\rangle + \langle r_j^+ s_i\rangle. \tag{16}$$

Our main ansatz is that the m-particle correlations $\sigma_{j_1,\ldots,j_m}(t) = \langle \delta_{j_1}(t)\cdots\delta_{j_m}(t)\rangle$ with $\delta_j(t) = s_j(t) - \langle n_j(t)\rangle$ scale with system size as

$$\sigma_{j_1,\ldots,j_m}(t) = O(N^{-m/2}), \quad \text{for } j_k \neq j_l. \tag{17}$$

We first show how the ansatz (17) allows to close the system (15, 16).
We assume that functional dependence of the rates on the sate variables is of the form $f(s_1/N,\ldots,s_N/N)$. This includes, for example, rates of the form $f(\sum\lambda_k s_k/N)$ like the ones used in the examples analyzed. We further assume that the rates can be expanded as a power series:

$$f(s_1/N,\ldots,s_N/N)=a_0+\sum_{i_1=1}^{N}a_{i_1}\frac{s_{i_1}}{N}+\frac{1}{2!}\sum_{i_1,i_2=1}^{N}a_{i_1,i_2}\frac{s_{i_1}s_{i_2}}{N^2}+\cdots+\frac{1}{k!}\sum_{i_1,\ldots,i_k=1}^{N}a_{i_1,\ldots,i_k}\frac{s_{i_1}\cdots s_{i_k}}{N^k}+\ldots \tag{18}$$

There are N^k terms in the k'th summand, $\sum_{i_1,\ldots,i_k=1}^{N}$, giving a total contribution of order $O(N^0)$. The terms in the right hand side of (15) are of the form:

$$\frac{\langle s_{i_1}\ldots s_{i_k}\rangle}{k!}=\frac{\langle(\delta_{i_1}+\langle s_{i_1}\rangle)\ldots(\delta_{i_k}+\langle s_{i_k}\rangle)\rangle}{k!}=\sum_{l=0}^{k}\frac{\delta^l\langle s\rangle^{k-l}}{l!(k-l)!}=\sum_{l=0}^{k}\frac{O(N^{-l/2})}{l!(k-l)!}, \tag{19}$$

where δ^l corresponds to a term of the form $\langle\delta_{j_1}(t)\cdots\delta_{j_l}(t)\rangle$, $\langle s\rangle^{k-l}$ corresponds to $\langle s_{i_1}\rangle\cdots\langle s_{i_{k-l}}\rangle$ and the last equality holds due to our ansatz. We see that the dominant terms are those with $l=0$, which correspond to products of mean values of the form $\langle s_{i_1}\rangle\cdots\langle s_{i_k}\rangle$. We conclude that the ansatz allows to do the substitution $\langle s_{i_1}\ldots s_{i_k}\rangle\rightarrow\langle s_{i_1}\rangle\cdots\langle s_{i_k}\rangle+O(N^{-1/2})$ in the evolution equations for the mean values.

The evolution equations for the correlations read:

$$\frac{d\sigma_{i,j}}{dt}=\langle(r_i^-+r_i^+)s_i\delta_j\rangle+\langle(r_j^-+r_j^+)s_j\delta_i\rangle+\langle r_i^+\delta_j\rangle+\langle r_i^+\delta_j\rangle. \tag{20}$$

In this case, the terms are of the form $\langle s_{i_1}\ldots s_{i_k}\delta_r\rangle=\langle(\delta_{i_1}+\langle s_{i_1}\rangle)\ldots(\delta_{i_k}+\langle s_{i_k}\rangle)\delta_r\rangle$ with $r=i,j$. Due to the presence of δ_s, the term in which only averages appears vanishes. Reasoning as before, we see that the dominant terms are those proportional to $\sigma_{i_l,s}$, while those proportional to higher-order correlations can be neglected. In this case, the ansatz allows to do the substitution $\langle s_{i_1}\ldots s_{i_k}\delta_r\rangle\rightarrow\langle s_{i_1}\rangle\cdots\langle s_{i_k}\rangle\sum_{l=1}^{k}\frac{\sigma_{i_r}}{\langle s_{i_r}\rangle}+O(N^{-3/2})$. In this way, the evolution equation for the correlations depend, at first order, only on averages and correlations and not on higher order moments.

The validity of the ansatz (17) itself can be established a posteriori by checking that the results obtained using the ansatz are consistent with it. In this section, we will link its validity with the well-known van Kampen's ansatz [1] that is the basis for the systematic system-size expansion.

Van Kampen's ansatz consists on assuming that the variable of interest has a deterministic part of order Ω plus a stochastic part of order $\Omega^{1/2}$, i.e. $n = \Omega\phi(t) + \Omega^{1/2}\xi$, where Ω is a parameter of the system that controls the relative size of the changes due to elementary processes, typically the system size.

In our system the role of the parameter Ω is played by the total number of particles N. As briefly stated in the main text, we cannot expect that the single-particle variables that we are considering obey van Kampen's ansatz, since they are not extensive. Our variables $s_i = 0, 1$ have a deterministic and stochastic part that are both of order zero respect to N (note that $\sigma^2[s_i] = \langle s_i\rangle(1 - \langle s_i\rangle)$). However, the macroscopic variable $n = \sum s_i$ is indeed extensive and we can expect that it will follow van Kampen's ansatz: $n = N\phi(t) + N^{1/2}\xi$. This implies that the m-th central moment of n will scale as $N^{m/2}$, i.e:

$$\langle (n - \langle n\rangle)^m\rangle = \sum_{j_1,\ldots,j_m} \sigma_{j_1,\ldots,j_m} = O(N^{m/2}). \tag{21}$$

Now, assuming that $\sigma_{j_1,\ldots,j_m} = f_m(N)\tilde{\sigma}_{j_1,\ldots,j_m}$ for $j_k \neq j_l$, with $\tilde{\sigma}_{j_1,\ldots,j_m}$ independent of N i.e. the m-particle correlations are all or the same order in N, so that $\sum_{j_1\neq j_2\neq,\ldots,\neq j_m} \tilde{\sigma}_{j_1,\ldots,j_m}$ scales as N^m (note that there are of the order of N^m terms in the sum), we obtain our main ansatz, $\sigma_{j_1,\ldots,j_m} = O(N^{-m/2})$ for $j_k \neq j_l$. We have only considered terms with $j_k \neq j_l$ in the sum (21); terms with repeated sub-indexes can be expressed as lower order ones. For example, if the index j_1 is present k times, and the others are all different, we find:

$$\begin{aligned}
&\sigma_{j_1,j_1,\ldots,j_1,j_2,\ldots j_{m-k+1}} = \langle (s_{j_1} - \langle s_{j_1}\rangle)^k \delta_{j_2}\ldots\delta_{j_{k-k+1}}\rangle \\
&= \sigma_{j_2,\ldots j_{m-k+1}}(-\langle s_{j_1}\rangle)^k + \langle \delta_{j_2}\ldots\delta_{j_{m-k+1}} \sum_{i=0}^{k-1} \binom{k}{i} (-\langle s_{j_1}\rangle)^i s_{j_1}\rangle \\
&= \sigma_{j_2,\ldots j_{m-k+1}}[(1 - \langle s_{j_1}\rangle)^k\langle s_{j_1}\rangle + (1 - \langle s_{j_1}\rangle)(-\langle s_{j_1}\rangle)^k] + \sigma_{j_1,\ldots,j_{m-k+1}}[(1 - \langle s_{j_1}\rangle)^k - (-\langle s_{j_1}\rangle)^k]
\end{aligned} \tag{22}$$

as can be see expanding $(s_{j_1} - \langle s_{j_1}\rangle)^k$ and keeping in mind that $s_i^2 = s_i$. The number of such terms in the sum (21) is $O(N^{m-k+1})$, so they give smaller contribution that terms with all sub-indexes different. Proceeding order by order from $k = 1$, we see that our main ansatz (17) follows from (21).

We point out that in systems of heterogeneous particles we do not have a closed description for the global, extensive, variable n so van Kampen's expansion cannot be used. Instead we derive the implications of van Kampen's ansatz over the correlations of the microscopic variables. (17) is a simple and convenient expression that in general allows to close the equation

for the moments (15,16). Often, however it is not necessary, and a weaker condition of the form (21), that directly follows from van Kampen's ansatz without further assumptions, is sufficient.

Van Kampen's ansatz is generally valid when the macroscopic equations have a single attracting fixed point, when the system displays small fluctuations around the macroscopic state. The general method explained here is expected to be valid under similar conditions. An interesting topic for future research will be whether a system that has a single attracting fixed point in the absence of diversity always maintains this globally stable state when diversity is present, and whether a system that does not posses this globally stable fixed point can acquire it when diversity is added.

D. Details of the calculation for the Kirman model

In the Kirman model with distributed influence, the averages and correlations obey:

$$\frac{d\langle s_i\rangle}{dt} = \epsilon - (2\epsilon + \overline{\lambda})\langle s_i\rangle + N^{-1}\sum_k \lambda_k \langle s_k\rangle, \tag{23}$$

$$\begin{aligned}\frac{d\sigma_{i,j}}{dt} &= -2(2\epsilon+\overline{\lambda})\sigma_{i,j} + N^{-1}\sum_k \lambda_k\left(\sigma_{i,k}+\sigma_{j,k}\right) \\ &+ \delta_{i,j}\left[\epsilon + a + (\overline{\lambda}-2a)\langle s_i\rangle - 2\sum_k \frac{\lambda_k\sigma_{i,k}}{N}\right]\end{aligned} \tag{24}$$

with $a \equiv \sum_k \frac{\lambda_k\langle n_k\rangle}{N}$. Note that, due to the particular form of the rates, these equations are indeed closed. The first equation leads to a steady state value $\langle s_i\rangle_{\rm st} = \frac{1}{2}$, which implies $\langle n\rangle_{\rm st} = \frac{N}{2}$ (a property that comes from the symmetry $0 \leftrightarrow 1$). (24) is a linear system of equations for the correlations. The steady state correlations can always be obtained by inverting the matrix that gives the couplings. Obtaining a closed expression for $\sigma^2[n]$ in terms of the moments of λ is, however, not straightforward. From (24), we see that in the steady state:

$$\sigma_{i,j} = \frac{\sum_k \lambda_k \frac{\sigma_{i,k}+\sigma_{j,k}}{N} + \delta_{i,j}\left[\epsilon + \overline{\lambda}/2 - 2\sum_k \frac{\lambda_k\sigma_{i,k}}{N}\right]}{2(2\epsilon+\overline{\lambda})}, \tag{25}$$

from where we obtain

$$\sigma^2[n] = \sum_{i,j}\sigma_{i,j} = \frac{N(\epsilon+\overline{\lambda}/2) + 2C(1-1/N)}{2(2\epsilon+\overline{\lambda})}, \tag{26}$$

with $C \equiv \sum_{i,j} \lambda_j \sigma_{i,j}$. Multiplying (25) by λ_j and summing over j, one obtains:

$$C \;=\; \frac{d(1-2/N) + (\epsilon + \overline{\lambda}/2)\overline{\lambda} N}{4\epsilon + \overline{\lambda}}, \tag{27}$$

where $d \equiv \sum_{i,j} \lambda_i \lambda_j \sigma_{i,j}$. This is obtained again multiplying (25) by $\lambda_i \lambda_j$ and summing over i, j:

$$d \;=\; \frac{(\epsilon + \overline{\lambda}/2)\langle \lambda^2 \rangle N - 2e/N}{4\epsilon}, \tag{28}$$

where $e \equiv \sum_{i,j} \lambda_i^2 \lambda_j \sigma_{i,j}$. Using the ansatz $\sigma_{i,j} = O(N^{-1})$ we see that the last term of (28) is $O(N^0)$ (while the other are of $O(N)$), so to the first order we obtain :

$$\sigma^2_{\rm st}[n] = \frac{N}{4}\left[1 + \frac{\overline{\lambda}}{2\epsilon} + \frac{\sigma^2[\lambda]}{2\epsilon\left(4\epsilon + \overline{\lambda}\right)}\right] + O(N^0), \tag{29}$$

with $\sigma^2[\lambda] = \overline{\lambda^2} - \overline{\lambda}^2$. We have seen how the application of the ansatz (17) allows one to obtain closed expression for the global average and variance. Interestingly, in this particular example, it is possible to include all higher order terms to obtain an exact expression for d (which gives the exact expression for $\sigma^2[n]$ trough (27,26)), details are given in the appendix:

$$d = \frac{N(\epsilon + \overline{\lambda}/2)\sum_{k=0}^{\infty}\left(\frac{-2}{N(4\epsilon+\overline{\lambda})}\right)^k \overline{\lambda^{2+k}}}{4\epsilon + \overline{\lambda} - \sum_{k=0}^{\infty}\left(\frac{-2}{N(4\epsilon+\overline{\lambda})}\right)^k \overline{\lambda^{1+k}}} = \frac{N(\epsilon + \overline{\lambda}/2)\overline{\frac{\lambda^2}{1+\frac{2\lambda}{N(4\epsilon+\overline{\lambda})}}}}{4\epsilon + \overline{\frac{2\lambda^2}{N(4\epsilon+\overline{\lambda})+2\lambda}}} \tag{30}$$

The second equality holds as long as [2] $\lim_{m\to\infty} \overline{\frac{\lambda^{m+2}}{1+\frac{2\lambda}{N(4\epsilon+\overline{\lambda})}}}\left(\frac{2}{N(4\epsilon+\overline{\lambda})}\right)^m = 0$. A sufficient condition for this is $\lambda_i < \frac{(\overline{\lambda}+4\epsilon)N}{2}$, $\forall i = 1, \ldots, N$. When the λ_i's are i.i.d. random variables, the probability that this condition is satisfied approaches one as N grows. This condition is actually necessary and sufficient for the first equality in (30) to hold (see appendix).
We finally obtain the following exact expression for the variance:

$$\sigma^2_{\rm st}[n] = \frac{N}{4}\left[1 + \frac{2\overline{\lambda}(1-1/N)}{4\epsilon + \overline{\lambda}} + (N - 3 + 2/N)\frac{\overline{\frac{\lambda^2}{N(4\epsilon+\overline{\lambda})+2\lambda}}}{2\epsilon + \overline{\frac{\lambda^2}{N(4\epsilon+\overline{\lambda})+2\lambda}}}\right] \tag{31}$$

We see from (31) that higher order corrections to $\sigma^2[n]$ depend on higher order moments of the distribution of λ over the population.

Expressions (29, 31) refer to the variance of n in a population with given values for the parameters of each agent, λ_i, so the averages are population averages i.e. $\overline{g(\lambda)} = \sum_{i=1}^{N} g(\lambda_i)/N$.

In the case that the parameters of the agents are random variables, the population averages themselves, $\overline{g(\lambda)}$, become random variables. To compute the expected (average) value of (29, 31), $\widehat{\sigma^2[n]}$, one has to average over the distribution of $\overline{g(\lambda)}$, which depends on the distribution $f(\lambda)$ of the $\lambda_i's$ (we are assuming $\lambda_i's$ i.i.d. random variables). This averages were obtained numerically, by evaluating expressions (29, 31) over the same realizations of the λ_i's that were used in the numerical simulations. One can use the approximation $\widehat{\overline{g(\lambda)}} \simeq \widehat{g(\lambda)}$, that works better the larger the N and the lower the variance σ^2_λ, and that, due to the law of large numbers, is valid in the limit $N \to \infty$. In Fig.2 of the main text we compare the average of the analytical expression (31) with results coming from numerical simulations. We find perfect agreement and see that at first order the dependence of $\sigma^2[n]$ with $\sigma^2_\lambda \equiv \widehat{\lambda^2} - \widehat{\lambda}^2$ is linear and independent of the form of the distribution, as indicated by (29). Higher order corrections are noticeable for higher levels of diversity.

In the case of heterogeneity in the preference of the agents for the states, as indicated in the main text, the variance is given by:

$$\sigma^2[n]_{st} \;=\; \frac{N}{4(\epsilon + \frac{2\lambda}{N})}\left[\overline{\epsilon^+}(1+\frac{\lambda}{\epsilon}) - \frac{\overline{\epsilon^+}^2}{\epsilon}\left(\frac{\lambda}{2\epsilon}+1\right) - 2\frac{\sigma^2[\epsilon]}{2\epsilon+\lambda}\right], \tag{32}$$

In this case, the average of (32) over the distribution of parameters can be easily computed, giving:

$$\begin{aligned}\widehat{\sigma^2[n]}_{st} \;=\; \frac{N}{4(\epsilon + \frac{2\lambda}{N})}\Bigg[&\widehat{\epsilon^+}\left(2+\frac{\lambda}{\epsilon}\right) - \widehat{\epsilon^+}^2\left(\frac{\lambda}{2\epsilon^2}+\frac{1}{\epsilon}\right) \\ &-\sigma^2_{\epsilon^+}\left(\frac{2\epsilon+\lambda/N}{\epsilon(2\epsilon+\lambda)} + \frac{\lambda}{2\epsilon^2 N}\right)\Bigg],\end{aligned} \tag{33}$$

The correlation function can be derived as follows (we exemplify the derivation in the case of distributed influence, for other types of heterogeneity, the derivation is similar):
(23) is an equation for the conditional averages $\langle s_i|\{s_l(t_0)\}\rangle$ if we set $\{s_l(t_0)\}$ as initial conditions. It implies:

$$\frac{da}{dt} = \epsilon\lambda - 2\epsilon a \to a(t_0+t) = \frac{\lambda}{2}(1-e^{-2\epsilon t}) + a(t_0)e^{-2\epsilon t}, \tag{34}$$

with $a \equiv \sum_k \lambda_k \langle s_k|\{s_l(t_0)\}\rangle/N$. Noticing that (23) is equal to $\frac{d\langle s_i\rangle}{dt} = \epsilon - (2\epsilon+\lambda)\langle s_i\rangle + a(t)$,

we obtain:

$$\langle s_i(t_0+t)|\{s_k(t_0)\}\rangle = \frac{1}{2}(1-e^{-(2\epsilon+\overline{\lambda})t}) + \frac{a(t_0)-\overline{\lambda}/2}{\overline{\lambda}}e^{-2\epsilon t}(1-e^{-\overline{\lambda}t}) + s_i(t_0)e^{-(2\epsilon+\overline{\lambda})t}. \quad (35)$$

Using now $K_{st}[n](t) = \langle\langle n(t_0+t)|n(t_0)\rangle n(t_0)\rangle_{st} - \langle n\rangle_{st}^2 = \sum_{i,j}\langle\langle s_i(t_0+t)|\{s_k(t_0)\}\rangle s_j(t_0)\rangle - \frac{N^2}{4}$ (remember $\langle n\rangle_{st} = N/2$), and after some straightforward algebra, we obtain:

$$K_{st}[n](t) = (\sigma_{st}^2 - C/\overline{\lambda})e^{-(2\epsilon+\overline{\lambda})t} + C/\overline{\lambda}e^{-2\epsilon t}. \quad (36)$$

From (25) we get $C/\overline{\lambda} = \frac{2\epsilon+\overline{\lambda}}{2\overline{\lambda}(1-1/N)}(\sigma_{st}^2 - N/4) \equiv u$, showing that (36) is equal to the expression displayed in the main text.

E. Appendix

We start with equation (25):

$$\sigma_{i,j} = \frac{\sum_k \lambda_k \frac{\sigma_{i,k}+\sigma_{j,k}}{N} + \delta_{i,j}\left[\epsilon + \overline{\lambda}/2 - 2\sum_k \frac{\lambda_k\sigma_{i,k}}{N}\right]}{2(2\epsilon+\overline{\lambda})}. \quad (37)$$

Using the rescaled variables $\tilde{\sigma}_{i,j} \equiv 4\sigma_{i,j}, \tilde{\lambda}_k \equiv \frac{\lambda_k}{2(2\epsilon+\overline{\lambda})N}$, and defining $S_n := \sum_{i,j=0}^{N} \tilde{\lambda_i}^n \tilde{\lambda}_j \tilde{\sigma}_{i,j}$, we obtain:

$$S_{n+1} = \frac{N\overline{\tilde{\lambda}}-1}{2}S_n + \frac{N}{2}\left(\overline{\tilde{\lambda}^n}S_1 + \overline{\tilde{\lambda}^{n+1}}\right). \quad (38)$$

Defining now $G_n := \left(\frac{2}{N\overline{\tilde{\lambda}}-1}\right)^n S_n, T_M := \sum_{n=1}^{M} G_n$, we arrive to:

$$G_{n+1} = G_n + \left(\frac{2}{N\overline{\tilde{\lambda}}-1}\right)^{n+1}\frac{N}{2}\left[G_1\left(-\frac{\overline{\lambda}+4\epsilon}{4(2\epsilon+\overline{\lambda})}\right)\overline{\tilde{\lambda}^n} + \overline{\tilde{\lambda}^{n+1}}\right], \quad (39)$$

$$T_{M+1} - G_1 = T_M + \frac{N}{2}\sum_{n=1}^{M}\left[\left(\frac{2}{N\overline{\tilde{\lambda}}-1}\right)^n\left(\frac{2\overline{\tilde{\lambda}^{n+1}}}{N\overline{\tilde{\lambda}}-1} + G_1\overline{\tilde{\lambda}^n}\right)\right]. \quad (40)$$

If $\lim_{M\to\infty} G_M = 0$, we see that:

$$G_1 = -\frac{\frac{N}{2}\sum_{n=1}^{\infty}\left(\frac{2}{N\overline{\tilde{\lambda}}-1}\right)^{n+1}\overline{\tilde{\lambda}^{n+1}}}{1+\frac{N}{2}\sum_{n=1}^{\infty}\left(\frac{2}{N\overline{\tilde{\lambda}}-1}\right)^n\overline{\tilde{\lambda}^n}}. \quad (41)$$

Going back to the original variables, we finally obtain, with the notation of the main text:

$$d = \frac{\frac{N^3(\epsilon+\overline{\lambda}/2)(4\epsilon+\overline{\lambda})}{4}\sum_{n=1}^{\infty}\left(\frac{-2}{(\overline{\lambda}+4\epsilon)N}\right)^n\overline{\lambda^{n+1}}}{1+\frac{N}{2}\sum_{n=1}^{\infty}\left(\frac{-2}{(\overline{\lambda}+4\epsilon)N}\right)^n\overline{\lambda^n}}, \quad (42)$$

which can be rewritten in the form (30), completing the proof.

The condition of convergence is:

$$\lim_{M\to\infty} G_M = \lim_{M\to\infty} \sum_{i,j=1}^{N} \left(\frac{-2\lambda_i}{(\overline{\lambda}+4\epsilon)N} \right)^M \frac{2\lambda_j}{(2\epsilon+\overline{\lambda})N} \sigma_{i,j} = 0. \tag{43}$$

A necessary and sufficient condition for this is $\lambda_i < \frac{(\overline{\lambda}+4\epsilon)N}{2}, \forall i = 1, \dots N$. When the parameters λ_i are i.i.d. r. v. the probability of this typically approaches 1 as N grows.

---

[1] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, (North-Holland, Amsterdam, 2004).

[2] Note that $\sum_{k=0}^{M} a^k \overline{\lambda^{k+2}} = \overline{\lambda^2 \sum_{k=0}^{M} a^k \lambda^k} = \overline{\lambda^2 \frac{1-a^{k+1}\lambda^{k+1}}{1-a\lambda}}$