We thank the recommender and the reviewers for their careful reading of our manuscript. We appreciate their enthusiastic comments, which helped us to improve our manuscript. We have modified our manuscript according to their comments. In this revised version, the changes are highlighted in red color. Please find below our detailed answer.

Recommender.

Point 1 *In particular, I would also like clearer motivations concerning your competition step and equation (4). This choice for competition may be responsible for the quick saturation with K, and this effect (that the stochastic IBM sees faster saturation with K than the deterministic equations) might go against the message that demographic stochasticity could increase the role of K, thankfully illustrated in Fig 4a at least. There is no doubt that, for any reasonable continuous equation, one can construct an IBM that will reproduce its results exactly, so I suppose that the IBM came first (or was modelled after the experiments) and equation (4) strives to reproduce its results? Or are you trying to make the strongest possible case by ensuring that, even when limiting the influence of K as much as possible (giving it no role before N reaches it in the IBM, and removing it from the max growth rate in (4)), it still plays a role? The motivation for these choices should made be clearer.*

Our answer. Our objective is to understand the effect of the carrying capacity *per se,* meaning that the parameter K should only impact the maximum sustainable population size (or density), but not (or marginally) the per capita growth rate under favorable conditions (optimal population density/size). This is why we constructed the growth term (4) instead of the classical growth term (3). As illustrated in SI1, with this classical growth term, increasing K also increases the maximum per capita growth rate. Obtaining an effect of K on the speed under this conditions would be quite obvious. With the growth term (4), when K is increased, the maximum sustainable density (K) changes, but the maximum and mean per capita growth rates are preserved. In the same spirit, the IBM (and in particular the competition term) was constructed so that the maximum expected per capita growth rate remains independent of K (see also our answer to Reviewer 1, Point 3).

Following your suggestion, we modified the paragraph below the introduction of eq (4) (pages 6-7). We also modified the paragraph describing the competition step in the IBM (page 8).

Point 2 *As it turns out, the two reviewers and myself unfortunately share a theoretical bent, but the experimental and statistical aspects appear to me to be sound. However, I think the paragraph on the experimental results could be put in more relief (perhaps given its own section).*

First, it is not so clear in the text why you would interpret this result as density-dependent dispersal, rather than any of your other mechanisms (hence putting it together with density-dependent dispersal results is not so obvious).

Second, you could perhaps present the qualitative results first, so as to limit the necessity to go back to the Methods to interpret what is said. For the less statistically minded, as I am, describing the effect of carrying capacity as an interaction between generation and modality is a counter-intuitive phrasing of a very intuitive idea.

Our answer. Previous results on the same population of Trichogramma highlighted the presence of density-dependent dispersal and the absence of an Allee effect (Morel-Journel et al. 2016). We also conducted a new ABC analysis on our own data to confirm the presence of density-dependent dispersal (details in Supplementary Information S3).

We modified the text in the Results section to make this clearer (pages 15-16).

Reviewer 1.

Point 1 *In the RD case (e.g. eq. 1) I wonder if there isn't a problem with the front speed being set by the low populations since with the assumption of diffusion the tail is infinitely long (i.e. no finite support)?*

Our answer. It is true that diffusion induces an infinite speed of propagation of the support of the solution, e.g., due to the strong parabolic maximum principle. However, the level lines of the solution propagate with a finite speed v. In the Fisher-KPP case, this speed depends on the growth function $f(u)$ only through its linearization around $u = 0$, i.e., through $f'(0)$. This means that the growth of populations with intermediate densities ($u > \epsilon > 0$) has no effect on the propagation speed. In other terms, the propagation is due to the low populations.

Another way to see that the speed v is set by low populations (in the Fisher-KPP case) is to observe that, for initial conditions of the form $u_0(x) = u(0, x) = e^{-\lambda x}$ for large x, the speed only depends on the tail of u_0: $v = \lambda D + f'(0)/\lambda$ when $0 < \lambda < \sqrt{f'(0)/D}$ and $v = 2\sqrt{f'(0)D}$ otherwise [see e.g., Uchiyama 1978]. In this work, we only focus on compactly supported initial conditions, which always lead to $v = 2\sqrt{f'(0)D}$.

We modified a paragraph on the Fisher-KPP model in the Introduction Section (page 3).

Point 2 *I think it is worth noting that equations 4 and 5 are not consistent in their units. This should probably be mentioned or amended.*

Our answer. It is true that the units in Sections 2.1 and 2.2 are different. We added a sentence in Section 2.2 (page 7).

Point 3 *The competition step for the IBM is quite strange. Is it perhaps the reason why the front is very sharp (fig. 2)? From a first glance, it appears that low populations are "safe" while beyond K they can collapse from this step. I realize that the reproduction step is the main cause of demographics and thus local population collapse, but I wonder what would happen to the front speed if the assumptions of the competition is changed. In particular, since the focus of this modeling approach is exactly to look at the importance of the carrying capacity, I think it would be good to make sure that there's nothing particular to this choice of implementation. Or perhaps there is a clear reason why this choice is sound from a mechanistic perspective?*

Our answer. The idea here is to check the effect of the maximum sustainable population size "*per se*". Thus, we consider an extreme case where the effect of competition is negligible when the population size is below the threshold K and then increases continuously. We believe that other assumptions would lead to a stronger effect of K on the propagation speed. Additionally, we note that, once the reproduction and dispersal steps have been defined, this is the only possible assumption on the competition step which lead to an expected population size K: the expected number of deaths due to competition must be $D_i - K$ when $D_i \geq K$, and is therefore equal to 0 when $D_i = K$ or takes lower values.

We modified the paragraph describing the competition step in the IBM (page 8).

Point 4 *Indeed the eq. 4 looks like a better "fit", but why? That is, it would be helpful to give intuition on what "went wrong" when constructing the model in eq. 3. Also where does the extra speed of IBMs come from?*

Our answer. In the IBM, the maximum expected per capita growth rate is $\max_{N_i(t)\geq 1} E[N_i\,(t+1)/N_i\,(t)] = \max_{N_i(t)\geq 1} Rg\,(N_i(t))/N_i\,(t) - E\,[\mu_i]/N_i\,(t)$, μ_i being the number of deaths due to competition. In all cases (with and without Allee effect), the maximum expected per capita growth rate is equal to R, and is therefore independent of K. In the case with Allee effect, this is consistent with eq (4), but not with the standard eq (3).

Regarding the extra-speed, as it both impacts the carrying capacity and the maximum per capita growth rate in eq (3), the parameter K leads to a stronger dependence of the speed with K, compared to the IBM. By “stronger dependence”, we mean that $v(K)$ increases faster. Here, the speeds have been standardized, so that the reaction-diffusion and IBM frameworks coincide when $K = 500$. Thus, as it increases faster, the reaction-diffusion curve corresponding to eq (3) is below the IBM and eq (4) curves for $K < 500$.

We added a paragraph in the MS (paragraph on strong Allee effect, IBM section, pages 8,9) and another paragraph in the Result section (strong Allee effect, page 15).

Point 5 *Would it not be possible to develop the IBM into a continuous form in the density dependence case? This should lead to a better explanation of the discrepancy between IBMs and RD in Fig. 4d*

Our answer. Unfortunately, we were not able to derive a formula for the speed in the RD case, with the density-dependent terms used in the IBM. An option would be to solve numerically the RD in this case.

Points 6 and 7 *Except for the third part of the discussion, which I find quite interesting, I find the discussion lacking as it mostly summarizes the results. I think it would be useful to add a discussion on the reasons for the different results shown, and to consider how one might distinguish between different scenarios where high K leads to faster speeds.*

Also, I feel that you could use the discussion to better give context on what your results mean from a practical perspective (i.e., what does this mean to conservation?). While I assume this paper is geared more to a theoretical crowd, some thoughts on implications would probably be useful to many people.

Our answer. We re-organized the discussion to address the reviewer’s critics. The first part is now focused on the direct interpretation of our results and the common characteristics shared by the three mechanisms. The second part discusses our results in the light of the pulled/pushed concepts, and their implications for the management of expanding populations, be they invasive or re-introduced.

Point 8 *In the SI, I am not sure if there is a clear added value in the content on the weak Allee effect, but if it is already there, why not compare the IBM results in Fig.3 (of the SI) to the theoretical ones mentioned earlier?*

Our answer. This was an oversight. Anyway, as suggested by the reviewer, we decided to remove the SI on weak Allee effects.

Reviewer 2.

Point 1. *Figure 5 presents the experimental results of "Mean number of colonized patches" versus "Generation". This Figure is not easily comparable to the theoretical results illustrated in Figure 4 (speed of the front versus carrying capacity). The authors should, at least, modified Figure 5 in order to show speed vs generation. If it is possible (maybe it is too complicated in experiments), the authors may show the experimental results of speed versus carrying capacity*

Our answer. As suggested by the reviewer, we added a figure (Fig. 5b) showing the speed vs generation.

Point 2. *Considering Equation (4), the parameter K is not only the carrying capacity (it is also related to the growth rate). Considering the Allee effect scenario, the authors should work with Equation (3), where K is the carrying capacity. These results are illustrated in Figure 4 (b and c) by dashed lines. These lines represent the formula v=(K-2*rho)*(rD/2K)^0.5, which also increases as the carrying capacity K increases. Considering this, a positive relation between speed and carrying capacity can be related to Allee effect or to positive density-dependent dispersal (K>2*rho). How do these results modify the conclusions?*

Our answer. Eq (4) has precisely been introduced in this work because K is related to the growth rate in the standard eq (3), whereas it is not the case in eq (4). This is also illustrated in Figs 1 and 2 of SI1: with the standard formula, the per capita growth rate strongly depends on K; with eq (4), K changes the value of the positive stable steady state, while keeping constant the maximum value of the per capita growth rate (and also its mean value, as mentioned in the main text).

Regarding the fact that a positive relationship between v and K can be related to an Allee effect or to positive density-dependent dispersal (or other mechanisms leading to pushed waves), this is acknowledged in the Discussion Section (part: "Carrying capacity impacting propagation speed, a proxy for pulled/pushed propagation?")

Point 3. *Why do the authors consider d^2(u^2)/dx^2 and not u*d^2(u)/dx^2 (as it is considered in Lejeune et al. 2002, Physical Review E 66, 010901) on the positive density-dependent dispersion? On the discussion about pulled/pushed propagation, the authors suggest that positive density-dependent dispersion leads to a pushed propagation. Is this conclusion related with the specific density-dependent dispersion relation considered by the authors? Is it more general?*

Our answer. Considering the term $u\partial^2 u/\partial x^2$ instead of $\partial^2 u^2/\partial x^2$ would be inappropriate here. We explain why. If we consider a simple equation with density-dependent dispersal, but without birth and death, the total population $P(t) = \int u(t,s)ds$ should be conserved. However, it is not the case with $u\partial^2 u/\partial x^2$: integrating by parts (with respect to x) the equation $\partial_t u = u\partial^2 u/\partial x^2$, we get $P'(t) = -\int \left(\frac{\partial u}{\partial x}\right)^2 (t,s)ds < 0$. As a consequence, the total population size decays. Thus means that the dispersal term $u\partial^2 u/\partial x^2$ is not conservative, contrarily to $\partial^2 u^2/\partial x^2$, and is therefore not appropriate for our purpose.

Regarding the fact that positive density-dependent dispersal leads to a pushed propagation, we believe that this conclusion should be quite general. Mechanisms that lead to an advantage for the individuals situated in the bulk of the population (here, higher mobility), compared to the individuals situated at the edge of the wave, tend to lead to pushed propagations. But we did not prove this result here.

Point 4. *Consider other references may improve the paper.*

• *In addition to Turchin 1998 and Hastings et al. 2005, at the beginning of the third paragraph of the introduction, I suggest Gilad et al. 2004 (Physical Review Letters 93, 098105) and Gilad et al. 2007 (Journal of Theoretical Biology 244, 680-691).*
• *In addition to Turchin 1998 and Lewis & Kareiva 1993, at the first paragraph in 2.1 Strong Allee effect, I suggest Clerc et al. 2005 (Physical Review E 72, 056217).*

We added the reference *Gilad et al. 2007 (Journal of Theoretical Biology 244, 680-691).* The reference *Clerc et al. 2005 (Physical Review E 72, 056217)* indeed uses a growth function of type (3), but the carrying capacity is not mentioned explicitly (K=1), so we decided not to include it in the MS.