We thank all reviewers for carefully reading our manuscript and their valuable comments. Our response is as follows.

Response to Reviewer #1

>> *...it's not clear why this is important. If the brain wants to send information, it will probably use axons.*
We dispute "it will probably use axons" because adult brains cannot grow long-distance axons, which could only happen as an evolutionary process. If, as a consequence of learning, a brain wants to transfer information reliably between two areas that previously were not functionally connected, it can use the heterogeneous FFNs in this study. We will add this argument to the Discussion.

>> *seemed to be several fixed points in the (α, σ) plane . . . non-generic behavior . . . if there was fine tuning involved*
The mechanism requires no fine-tuning. The signal transformations by neighboring layers (Fig. 3E left) need not be in exactly opposite directions but can only be sufficiently different to prevent continuous flows (Fig. 3E middle, right).

Response to Reviewer #2

>> *However, it is not clear how general . . . inhibitory neurons play an essential role . . .*
We agree that inhibitory neurons are important. However, in more realistic contexts, the mechanism is mostly about how the initial part of the network input propagates before other mechanisms, such as recurrent inhibition, activate (see also Ref. [1]). An exception can be fast feedforward inhibition, so we tested different levels of feedforward inhibition, and found that the main results do not change as long as it is not too strong (e.g. Appendix Fig. A3).

>> *previous works . . . introduce heterogeneity into the network choosing parameters of neurons . . .*
Thanks for the suggestion. We will complement the discussion with this track of related works.

Response to Reviewer #3

>> *1. . . . Is it possible with an intermediate value of β . . . the network simply relay the input . . ..*
No. No matter what β_w is used, it fixes the signal transformation property of neurons. Repeating the transform through multiple layers will accumulate signal distortion and lead to information loss, as explained in Fig. 1 and related text. We simulated the AL network model with neurons of β_w =-10 mV (Fig. R1, left), and found that d' for the ORN input detection and information transfer are impaired compared to the original, heterogeneous model.

>> *2. If we reverse the order of differentiator and integrator, would the input be reliably transmitted as well? . . .*
Yes. The reversed AL network also shows the stable power amplification and good information transfer (Fig. R1A-C, right and D). In the deep network models, we made similar observations. However, d' for the ORN input detection is suboptimal (Fig. R1B; dots are lower than solid lines), because ORNs fire sparsely with strong differentiator characteristics (Ref. [17]). Integrators are better suited for ORN's postsynaptic cells.

>> *. . . have previous network models considered the heterogeneity on the intrinsic parameters of neurons?*
Some previous studies considered heterogeneous neuronal property, but not in a laminar-specific manner. We will complement related discussion.

>> *I am quite confused about the dynamics (Eq. 1)...*
Thanks for carefully checking the equations. Eq. 1 had a typo, and the second term of the right hand side should be $-g_K w(V - E_K)$, not $-g_K(V - E_K)$. The correct equation was used in our source code. Note that w_∞ is a function of V, used for the equation for w. We will fix the equations in the revision.

>> *How does Fig. 4A lead to $\cdots$ the proposed model could adopt both spike rate and spike timing coding . . . ?*
We meant that Fig. 4A corresponds to spike timing (latency), in addition to the previous results (e.g. Fig. 2D) related to rate coding. We will clarify it in revision.

>> *Fig. 2C and Lines 144-146 . . . I am wondering whether β_w could be adjusted into an intermediate value . . .*
See our response to the question 1 and Fig. R1.

>> *. . . provide a figure demonstrating how the network varies . . . when β_w changes*
We will happily add the figures requested by the reviewer in the next version.

>> *how sensitive the reliable transmission depends on the values of β_w in two consecutive layers.*
Please refer to our response to the second questions of Reviewer #1.

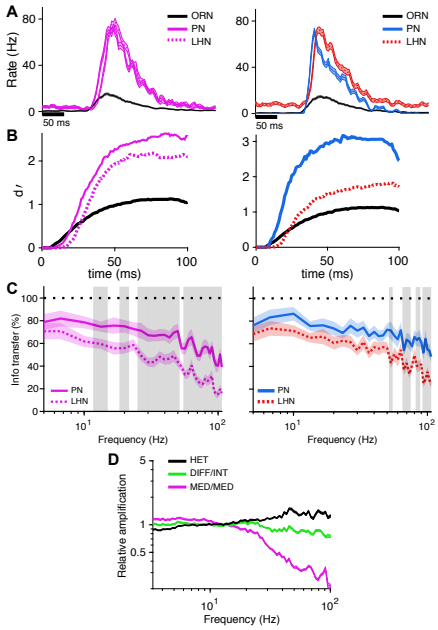

Figure R1: **A-C** Firing rates (A), d' (B), and information transfer (C) for the homogeneous AL network with β_w=-10 mV (left) and reversed heterogeneous model (right). **D** Power amplification of the original (black), reversed (green), and β_w = -10 mV network model. Blue: Differentiator, Red: Integrator, Magenta: β_w=-10 mV. Shade in C: P<0.01.