Reviewer 1: We thank the reviewer for the thorough review. We agree that our discussion of Seung et al. was not sufficient and we will address this in a future version. However, our contributions go beyond Seung et al.'s work in many aspects. We kindly ask the reviewer to reconsider the following contributions. (1) Our networks are spiking, which is different from the rate dynamics of Seung et al. We derived spiking dynamics as a greedy algorithm optimizing the minimax objective. (2) We made the connection between the minimax objective, and the tight and detailed balance observed in cortical networks by investigating the KKT conditions. The notion of balanced networks is not considered at all in Seung et al., but is an important phenomenon in neuroscience (Poo et al, 2009, Rupprecht et al, 2018). (3) Normative modeling (as in Hopfield networks) has been very fruitful in this field. We extend this tool to tightly and detailed balanced spiking neural networks of excitatory and inhibitory neurons. While our framework still suffers from some of the problems of Hopfield networks, it does extend the biological plausibility of normative modeling considerably. Using our framework, we constructed balanced spiking neural networks for solving various tasks (reconstruction, fixed point and manifold attractor dynamics) heavily studied in neuroscience. The E-I circuit architectures we provide are novel to the best of our knowledge. Such applications were not available in Seung et al. We indeed applied the results in Seung et al. as a tool to provide necessary conditions of convergence of the dynamics, however, the major contribution of our paper is to demonstrate that functional balanced E-I spiking networks can be designed from minimax objectives, which is conceptually novel and potentially useful for the neuroscience community.

Reviewer 2: We thank the reviewer for the enthusiastic support!

Clarification: (1) For the image reconstruction task, the number of excitatory neurons is $N_E = 400$. (2) For the fixed point attractor, we have $N_E = 15, N_I = 15$. For the specific example we show there are only 2 attractor states stored in this network. The attractors include specific inhibitory neurons and the network weights are pretrained given these specific E and I neurons. (3) To design weights for the attractor networks, we ask the combination of the connectivities (effective weight) $\mathbf{W}^{EE} - \mathbf{W}^{EI}\mathbf{W}^{II^{-1}}\mathbf{W}^{IE}$ to mimic the weight matrices in standard attractor models, in the examples that we show, we assume uniform inhibition for simplicity, such that the structure of the effective weight is contained in $\mathbf{W}^{EE}$, the matrices $\mathbf{W}^{EE}$, $\mathbf{W}^{EI}$ and $\mathbf{W}^{IE}$ are low rank. However, there are multiple ways of designing these matrices such that the effective weight remains the same. We will provide details in the appendix.

Minimax objectives: We thank the author for the inspiring question. (1) Dale's law is a structural constraint for biological plausibility, it is not yet clear to us whether Dale's law has any computational advantage. (2) We proposed minimax objective as a normative approach for designing spiking networks that obey Dale's law, and we have not focused on the advantage of minimax objective itself. However, minimax objectives can be a useful strategy when faced with uncertainty, the underlying intuition behind minimax is to look for the least-worst option out of all possible outcomes of an action. Previous works have demonstrated that minimax objectives can be used to solve problems such as source localization (Venkatesh et al, 2017) and sensorimotor control (Ueyama Y.,2014), and our methods could potentially be of use to design E-I spiking neural networks that accomplish these tasks. We will discuss these points.

Reviewer 3: We thank the reviewer for the feedback and we will revise our paper taking all his/her suggestions into account. We kindly request a reevaluation based on the following response to the weaknesses of our paper, which we believe arose from inclarity on our part. We also ask the reviewer to consider our response to the first reviewer for what we believe are our other main contributions than the identification of the minimax objective.

Minimax objective: Thank you for the great question which goes to the heart of our paper! 1) Absorbing the minus sign into the definition of inhibitory weights (W) in equation (2) does NOT lead to a minimizing dynamics on a quadratic function; it still is a minimax dynamics on equation (4) with weights redefined. To see why E-I circuit dynamics cannot be gradient descent, note that gradient descent on a minimization problem would give symmetric connections (defined as the weight times the $\pm$ sign in front in our convention). However, in networks that obey Dale's law, I-to-E connections have the opposite sign of E-to-I ones. 2) We agree that the difference between gradient descent-ascent on a minimax vs. gradient descent on a minimum objective is the dynamics converging to a saddle point vs. a minimum of an objective function. However, the dynamics of approaching a saddle point can be quite different than that of a minimum, although we do not discuss these aspects in our paper. Also note that both the saddle points and the minima can be fixed points of the optimizing dynamics, which can be either discrete or form a manifold as shown in our examples. We will expand our discussion of these points.

KKT conditions and spiking dynamics: We thank the reviewer for giving us an opportunity to clarify this potentially confusing point. The spiking dynamics derived in SI A performs a greedy optimization of the objective function. Due to the discontinuous spiking dynamics, the KKT conditions are never exactly satisfied, consistent with the reviewer's intuition about a biological circuit. The network constantly monitors for deviation from the KKT point, and corrects for it (imperfectly) by producing spikes when V_i is larger than the spiking threshold $V_{th}^i = \frac{1}{2}W_{ii}$. For the inactive (non-spiking) neurons, KKT conditions state that the voltage $V_i < 0$. This is smaller than a positive V_{th}^i, consistent with the neuron never spiking. We will expand our discussion of this point.

Variable substitution: We will clarify this. The "variable substitution trick" is merely the statement that the minimum of Eq.10 is equivalent to the saddle point of Eq.11. It holds for the optimum of the objective (or the steady state of the dynamics). We will also correct the typos the reviewer pointed out.