## Comments and differences of Philips et al.

The main difference between our approach and theirs is the formation of the preconditioner. Consider the block matrix:

$$\mathcal{K} = \begin{pmatrix} \mathcal{K}_{\mathrm{NS}} & \mathcal{K}_{\mathrm{C}}^T \\ -\mathcal{K}_{\mathrm{C}} & \mathcal{K}_{\mathrm{M}} \end{pmatrix},$$

where  $\mathcal{K}_{NS}$ ,  $\mathcal{K}_{C}$  and  $\mathcal{K}_{M}$  are the Navier-Stokes, coupling and Maxwell block systems, respectively. We took the approach to consider combining well-known preconditioners where as they went for an "all-at-once" block approach. The approaches are as follows:

$$\mathcal{P}_{\text{WGS}} = \begin{pmatrix} \mathcal{P}_{\text{NS}} & \mathcal{K}_{\text{C}}^T \\ -\mathcal{K}_{\text{C}} & \mathcal{P}_{\text{M}} \end{pmatrix},$$

we then combine the coupling terms in the Navier-Stokes preconditioner to form a block triangular preconditioner and

$$\mathcal{P}_{\mathrm{PSCEP}} = egin{pmatrix} \mathcal{X}_{\mathrm{NS}} & \mathcal{K}_{\mathrm{C}}^T \\ 0 & \mathcal{K}_{\mathrm{M}} \end{pmatrix} \quad \mathrm{where} \quad \mathcal{X}_{\mathrm{NS}} = \mathcal{K}_{\mathrm{NS}} + \mathcal{K}_{\mathrm{C}}^T \mathcal{K}_{\mathrm{M}}^{-1} \mathcal{K}_{\mathrm{C}}.$$

The resulting preconditioners have a similar block structure, however the approximations for  $\mathcal{K}_{\mathrm{M}}^{-1}$  and  $\mathcal{X}_{\mathrm{NS}}^{-1}$  are slightly different. The Maxwell preconditioned used is similar to the one we use, however is is triangular and uses a mass matrix instead of the Laplacian for the augmentation form. For the block Schur complement  $\mathcal{X}_{\mathrm{NS}}$ , the authors use a shifted LSC type approximation.

To check the performance of their preconditioner they perform a few numerical experiments whereas we do experiments as well as an eigenvalue analysis. I have coded the 3D cavity driven flow (the same conditions as they used) and have produced scalable results.

Finally, they use the same approximation for the coupling term that we do (4.10). Their arguments for using this are for high diffusion the  $\nabla \times \nabla \times$  becomes the dominant term. I quote "It can be seen that the operator (4.5) [the continuous form similar to (4.8) in our paper] should be approximately  $\text{Re}_m I$  in this case, as something close to the curl-curl operator is inverted while two curls are applied".