Appendix A

With two predictors (y1 and y2), and just looking at one indicator (x1) of a single latent variable, the covariance between x1 and y1 ($cov_{x1,y1}$) is

$$cov_{x1,y1} = \lambda_{x1}\beta_{y1}\sigma_{y1} + \lambda_{x1}\beta_{y2}\sigma_{y1,y2},$$

where λ_{x1} is the factor loading for x1, β_{y1} is the regression coefficient for y1, σ_{y1} is the variance of y1, and $\sigma_{y1,y2}$ is the covariance between y1 and y2.

This means that when predictor covariance is high, the estimation of the second regression coefficient plays a large role. Thus, whenever we over-penalize parameters, either because we don't have a large enough sample size to estimate them, or we just want sparsity, this bias is not just incurred in the regression, but will trickle down to the factor loadings. This is always a problem, as λ_{x1} will try and "make up for" the downward bias of β_{y1}, but it can be exacerbated to a large extent when there is covariance among predictors. Add in a large numbers of predictors and this makes the problem much worse.