Figure S1: Example 1: Relative error, formula (2.9) of the Main text, as function of the number of iterations for the Conjugate Gradient method (upper panel) and the Tikhonov regularization parameter $\lambda$ (lower panel). Note that the x-axis of the lower panel is reversed. Both optimal number of iterations and regularization parameter are indicated by $\times$. 
Figure S2: Boxplots of 250 independent \( \pi_0 \) estimates from simulated data based upon a ‘true’ standard normal density of effect sizes, using 7 different methods: convest (Langaas et al., 2005), qvalue(Storey, 2003), our proposed nncg method and the four methods proposed by Long et al. (2012). Each of the 250 simulated data sets involves two groups, with 5 independent samples each.
Figure S3: Boxplot of Hellinger distances between 250 estimated density of effect sizes, from simulated data based upon a ‘true’ standard normal density of effect sizes, using our proposed nnCg method and the four methods proposed by Long et al. (2012). Each of the 250 simulated data sets involves two groups, with 5 independent samples each.
Figure S4: Boxplots of 250 $\pi_0$ estimates, each from a simulation study with a bimodal density of effect sizes, using 7 different methods: convest (Langaas et al., 2005), qvalue (Storey, 2003), our proposed nncg method and the four methods proposed by Long et al. (2012). Each of the 250 simulated data sets involves two groups, with 5 independent samples each.
Figure S5: Boxplot of Hellinger distances between 250 estimated density of effect sizes with the 'true' bitriangular density of effect sizes, using our proposed nncg method and the four methods proposed by Long et al. (2012). Each of the 250 simulated data sets involves two groups, with 5 independent samples each.
Figure S6: Example 3: Estimated $\pi_0$, average power, normalization constant for a) full sized data indicated with horizontal dashed lines and b) for 25 subsets of half the size of the data indicated with box-plots, using the three methods A, B and C.
Figure S7: Example 5: Upper panel: Estimated density of effect sizes using estimation method A (solid), B (dashed) and C (dotted). Lower panel: estimate power curves for sample sizes up to $5 \times$ larger than the pilot data where a small region around zero is excluded from the density of effect sizes.