Proof. (1) Recall that for an optimal solution with $w \neq 0$ and $\eta > 0$, we have $\gamma = 0$, so by $(*_\gamma)$ we have the equations

$$\sum_{i=1}^{p} \lambda_i = \frac{K_m}{2} \quad \text{and} \quad \sum_{j=1}^{q} \mu_j = \frac{K_m}{2}.$$

The point u_i fails to achieve the margin iff $\lambda_i = K_s = K_m/(\nu(p+q))$, so if there are p_f such points then

$$\frac{K_m}{2} = \sum_{i=1}^{p} \lambda_i \geq \frac{K_m p_f}{\nu(p+q)},$$

so

$$p_f \leq \frac{\nu(p+q)}{2}.$$

A similar reasoning applies if v_j fails to achieve the margin δ with $\sum_{i=1}^{p} \lambda_i$ replaced by $\sum_{j=1}^{q} \mu_j$.

(2) A point u_i has margin at most δ iff $\lambda_i > 0$. If

$$I_{\lambda>0} = \{i \in \{1, \ldots, p\} \mid \lambda_i > 0\} \quad \text{and} \quad p_m = |I_{\lambda>0}|,$$

then

$$\frac{K_m}{2} = \sum_{i=1}^{p} \lambda_i = \sum_{i \in I_{\lambda>0}} \lambda_i,$$

and since $\lambda_i \leq K_s = K_m/(\nu(p+q))$, we have

$$\frac{K_m}{2} = \sum_{i \in I_{\lambda>0}} \lambda_i \leq \frac{K_m p_m}{\nu(p+q)},$$

which yields

$$p_m \geq \frac{\nu(p+q)}{2}.$$

A similar reasoning applies if a point v_j has margin at most δ. We already observed that (†) implies that $p_m \geq 1$ and $q_m \geq 1$.

(3) This follows immediately from (1). $\square$

Observe that $p_f = q_f = 0$ means that there are no points in the open slab containing the separating hyperplane, namely, the points u_i and the points v_j are separable. So if the points u_i and the points v_j are not separable, then we must pick ν such that $2/(p+q) \leq \nu \leq \min\{2p/(p+q), 2q/(p+q)\}$ for the method to succeed. Otherwise, the method is trying to produce a solution where $w = 0$ and $\eta = 0$, and it does not converge (γ is nonzero). Actually, Proposition 54.1 yields more accurate bounds on ν for the method to converge, namely

$$\max\left\{\frac{2p_f}{p+q}, \frac{2q_f}{p+q}\right\} \leq \nu \leq \min\left\{\frac{2p_m}{p+q}, \frac{2q_m}{p+q}\right\}.$$