Multi-pathway benchmark: regret vs masking rate