Imagine you have some fixed problem

$$\min_{x \in \mathcal{X}} f(x)$$

If you were to run uSCG on this you would compute:

$$x^{k+1} = x^k + \gamma \text{lmo}_{\mathcal{D}}(\nabla f(x^k))$$

where $\mathcal{D} = \{x \in \mathbb{R}^d : \|x\| \leq 1\}$ is some norm ball; without loss of generality we will assume it is of radius 1.

Then, adding *decoupled* weight decay to this algorithm will result in

$$x^{k+1} = (1 - \lambda)x^k + \gamma \text{lmo}_{\mathcal{D}}(\nabla f(x^k))$$

We can reinterpret this as

$$x^{k+1} = (1 - \lambda)x^k + \lambda \text{lmo}_{\frac{\gamma}{\lambda}\mathcal{D}}(\nabla f(x^k))$$

i.e., we are now requiring that the weights are constrained to stay in the ball $\frac{\gamma}{\lambda}\mathcal{D}$ of radius $\frac{\gamma}{\lambda}$.
If we add *coupled* weight decay we would get

$$x^{k+1} = (1 - \lambda\gamma)x^k + \gamma \text{lmo}_{\mathcal{D}}(\nabla f(x^k))$$

We can reinterpret this as

$$x^{k+1} = (1 - \lambda\gamma)x^k + \lambda\gamma \text{lmo}_{\frac{1}{\lambda}\mathcal{D}}(\nabla f(x^k))$$

i.e., we are now requiring that the weights are constrained to stay in the ball $\frac{1}{\lambda}\mathcal{D}$ of radius $\frac{1}{\lambda}$.

Both of these are are *different* than if we added Tikhonov regularization.