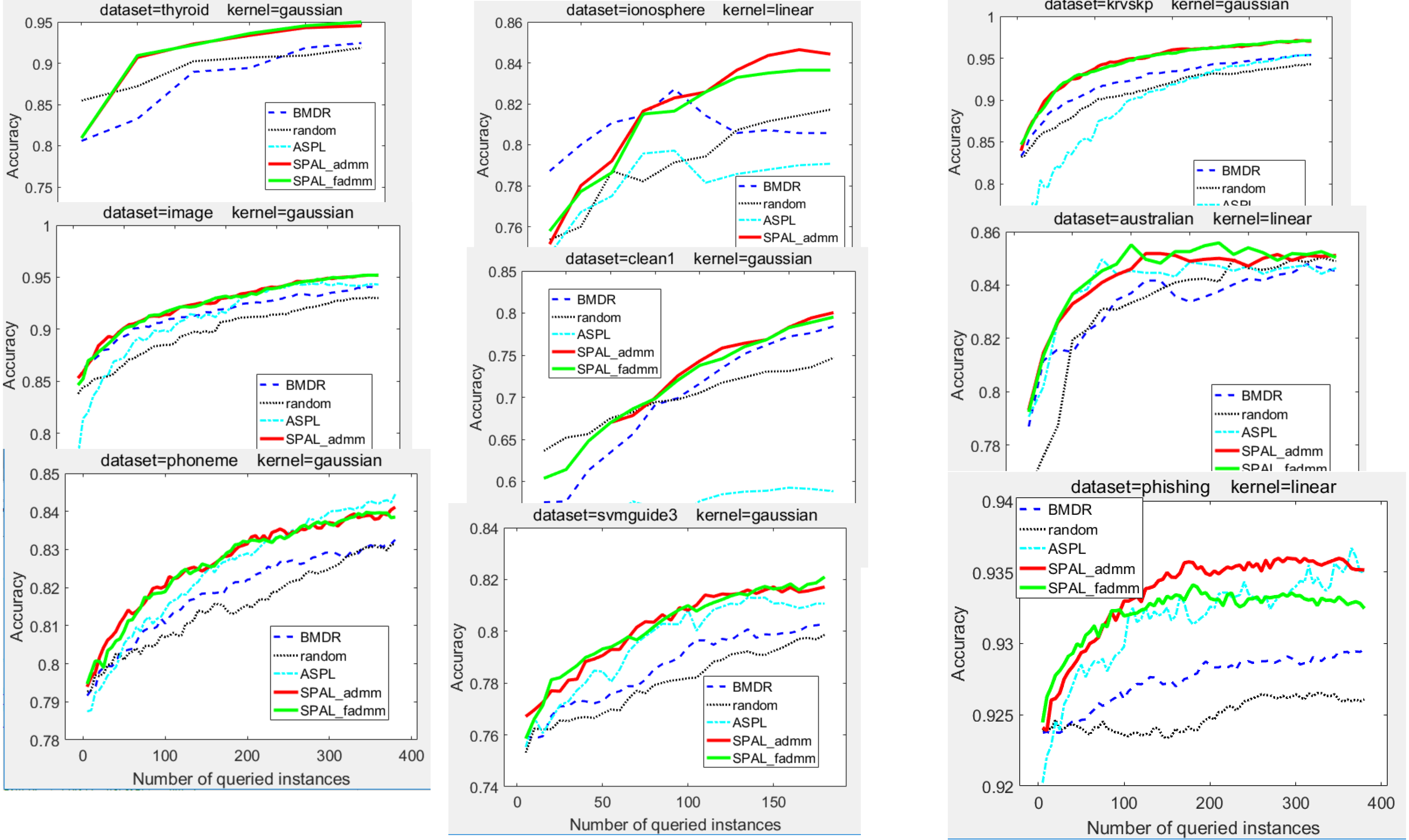
- 优化f时的判断**收敛精度**对性能影响很大，较小的精度一般性能会下降很多，也有极少数据集会变好
- 优化方法对性能影响较大
- 对label set的权去掉的话，性能在有些数据集变好有些数据集变差，有些变化不大
- 将SPL的权去掉的话，性能在大部分数据集上有所下降，小部分小的数据集有些上升
- 对参数λ的步长不太敏感，考虑测试一下初始值

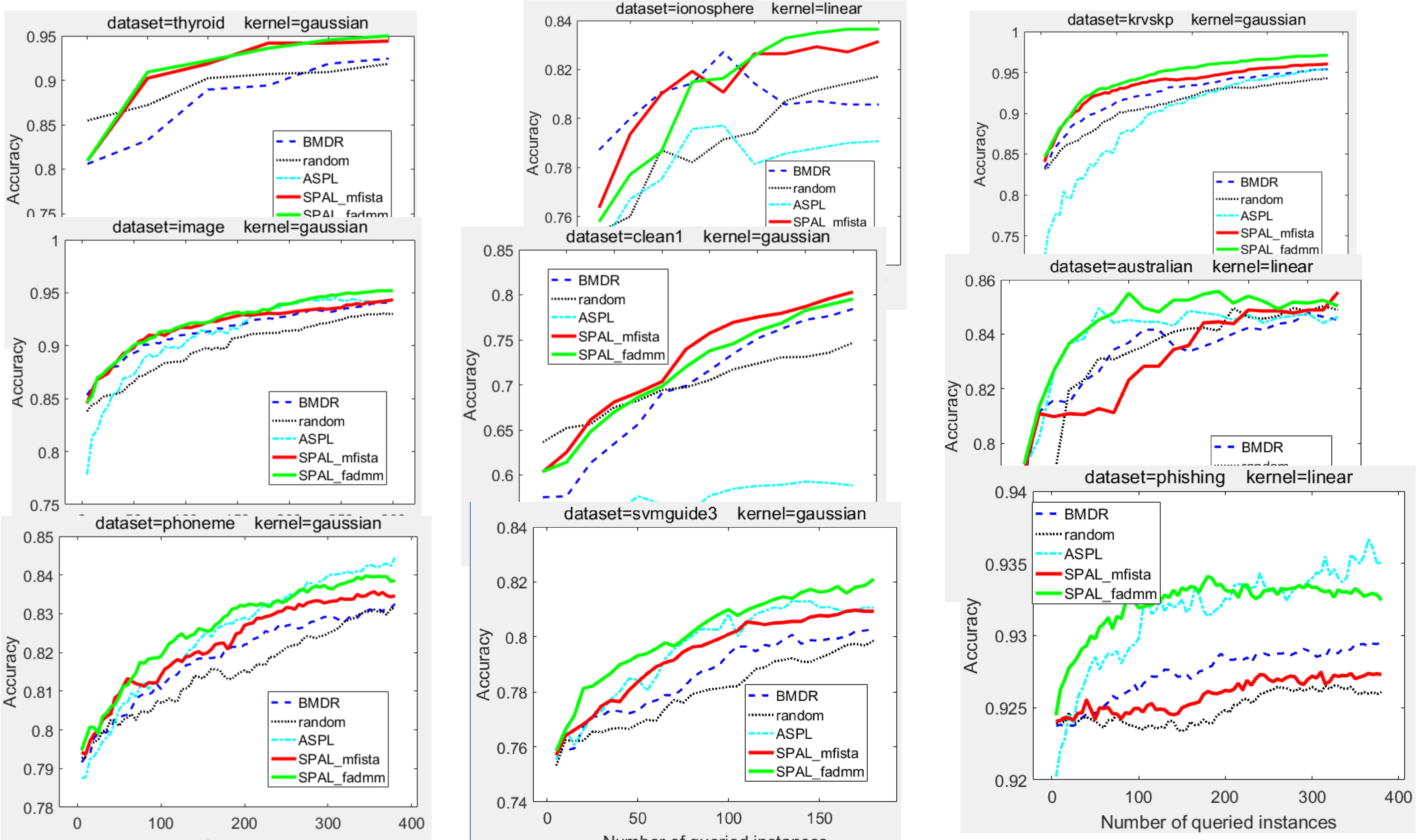
- ADMM ：对约束条件放松（红色）
- FADMM ：不放松约束，但过滤掉权较小的样本（绿色）

精度： 1e-4

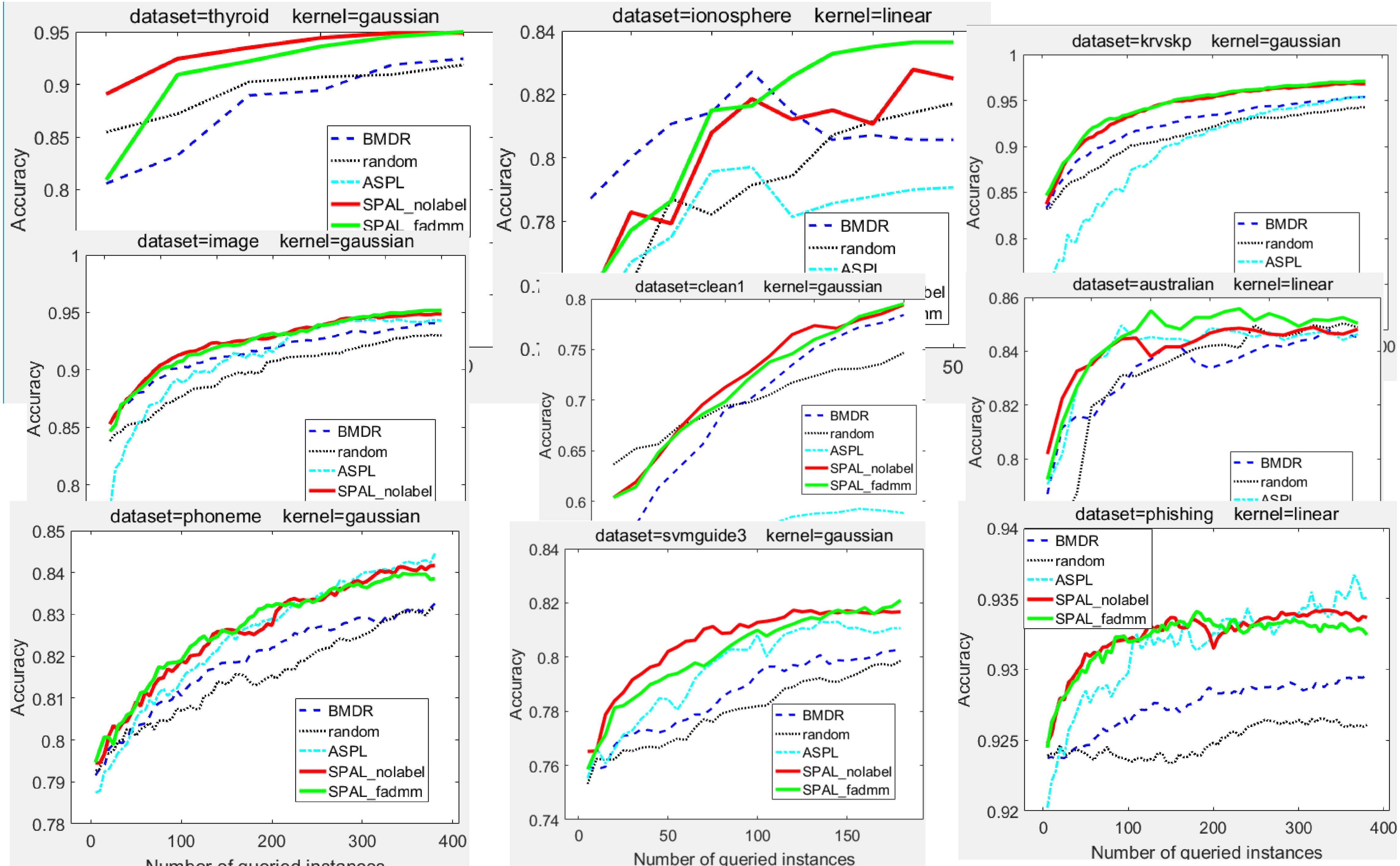


dataset=thyroid kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_admm
SPAL_fadmm
dataset=ionosphere kernel=linear
Accuracy
BMDR
random
ASPL
SPAL_admm
dataset=krvskp kernel=gaussian
Accuracy
BMDR
random
dataset=image kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_admm
dataset=clean1 kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_admm
SPAL_fadmm
dataset=australian kernel=linear
Accuracy
BMDR
random
ASPL
SPAL_admm
dataset=phoneme kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_admm
SPAL_fadmm
Number of queried instances
dataset=svmguide3 kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_admm
SPAL_fadmm
Number of queried instances
dataset=phishing kernel=linear
Accuracy
BMDR
random
ASPL
SPAL_admm
SPAL_fadmm
Number of queried instances

- MFISTA ：快速的近端梯度下降法
- FADMM ：不放松约束，但过滤掉权较小的样本（绿色）

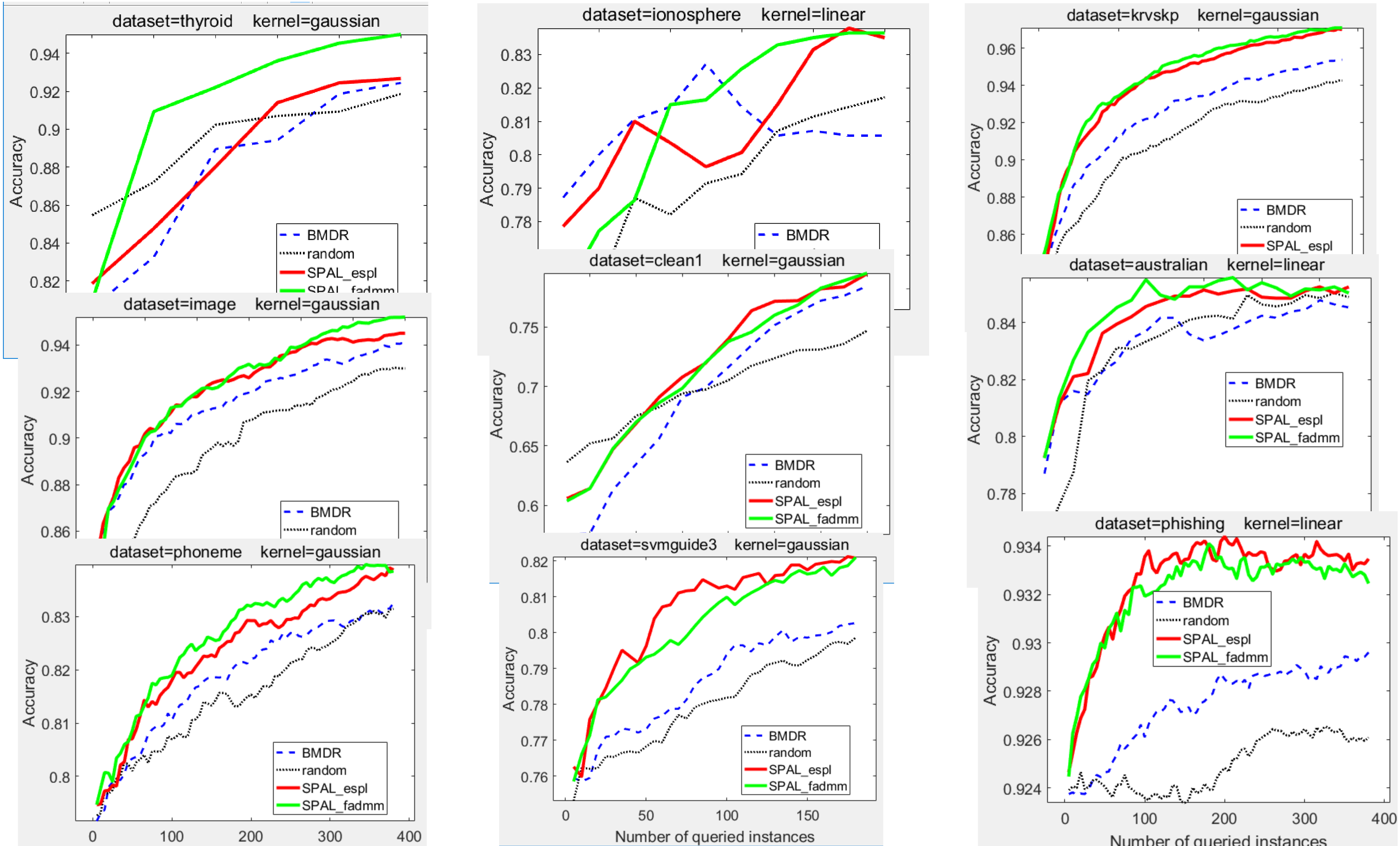

dataset=thyroid kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_mfista
SPAL_fadmm
dataset=ionosphere kernel=linear
Accuracy
BMDR
random
ASPL
SPAL_mfista
dataset=krvskp kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_mfista
SPAL_fadmm
dataset=image kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_mfista
SPAL_fadmm
dataset=clean1 kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_mfista
SPAL_fadmm
dataset=australian kernel=linear
Accuracy
BMDR
dataset=phoneme kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_mfista
SPAL_fadmm
dataset=svmguide3 kernel=gaussian
Accuracy
BMDR
random
ASPL
SPAL_mfista
SPAL_fadmm
dataset=phishing kernel=linear
Accuracy
BMDR
random
ASPL
SPAL_mfista
SPAL_fadmm
Number of queried instances

- NOLABEL:去掉label set的权（红色）
- FADMM ：不放松约束，但过滤掉权较小的样本（绿色）



dataset=thyroid kernel=gaussian
dataset=ionosphere kernel=linear
dataset=krvskp kernel=gaussian
dataset=image kernel=gaussian
dataset=clean1 kernel=gaussian
dataset=australian kernel=linear
dataset=phoneme kernel=gaussian
dataset=svmguide3 kernel=gaussian
dataset=phishing kernel=linear
Accuracy
Number of queried instances
BMDR
random
ASPL
SPAL_nolabel
SPAL_fadmm

- NOSPL	:去掉v(红色)
- FADMM	：不放松约束，但过滤掉权较小的样本（绿色）



dataset=thyroid kernel=gaussian
dataset=ionosphere kernel=linear
dataset=krvskp kernel=gaussian
dataset=image kernel=gaussian
dataset=clean1 kernel=gaussian
dataset=australian kernel=linear
dataset=phoneme kernel=gaussian
dataset=svmguide3 kernel=gaussian
dataset=phishing kernel=linear
Accuracy
Number of queried instances
BMDR
random
SPAL_espl
SPAL_fadmm

- QP：

$v_j \cdot w_j \left(\hat{y}_j - f(x_j)\right)^2$ where $\hat{y}_j = -sign(f(\boldsymbol{x}_j))$

We should have: $v_j \cdot w_j \ \left(f(\boldsymbol{x}_j)^2 + 2|f(\boldsymbol{x}_j)| + 1\right)$

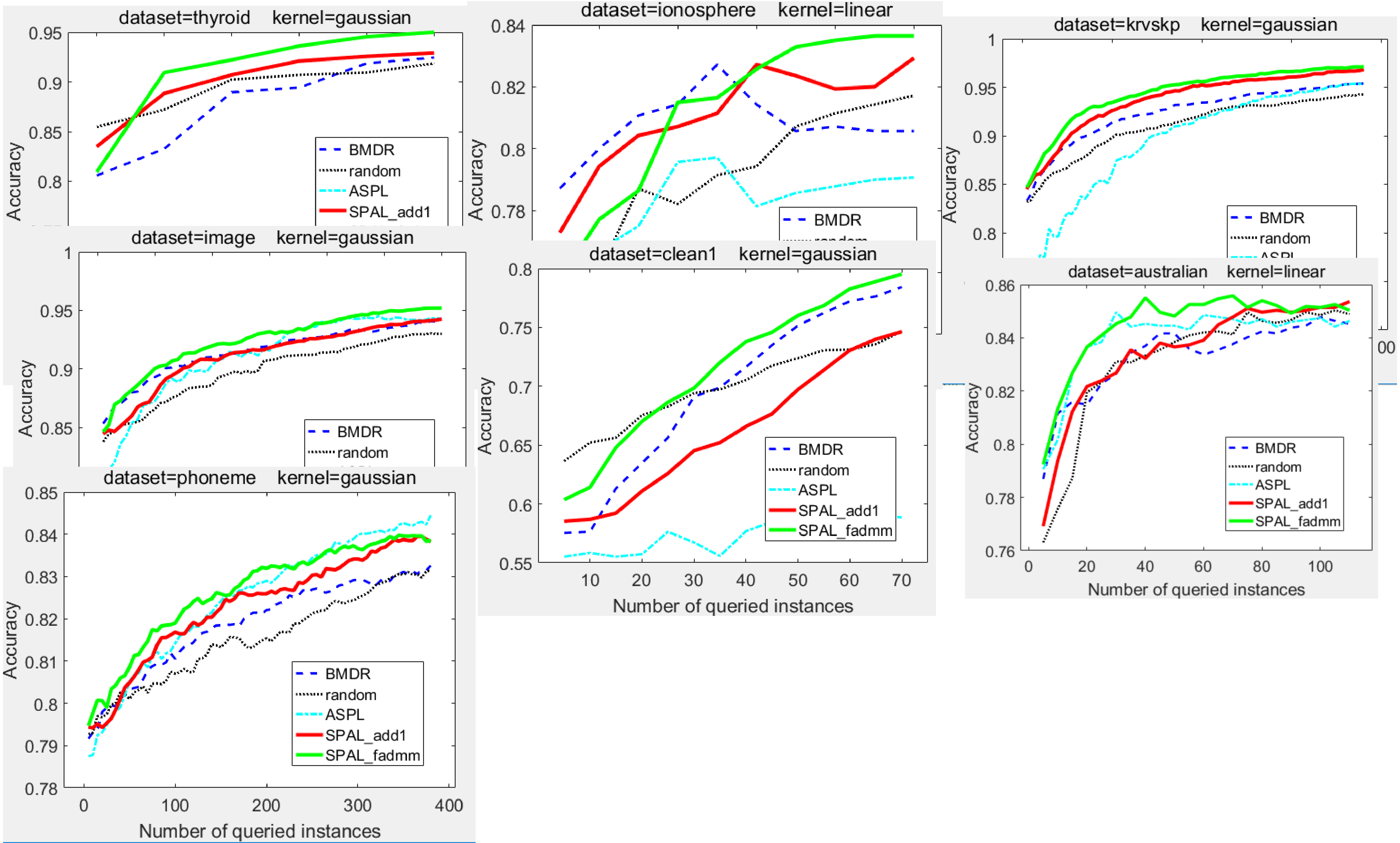
which can be rewritten as

$$\min_{\boldsymbol{\alpha}^T \mathbf{1}_u = b} \beta \boldsymbol{\alpha}^T K_1 \boldsymbol{\alpha} + (\beta \mathbf{k} + \mathbf{a})\boldsymbol{\alpha},$$

where $a_j = ||\mathbf{w}^T \phi(\mathbf{x}_j)||_2^2 + 2|\mathbf{w}^T \phi(\mathbf{x}_j)|$. This

BMDR KDD' 13

- ADD1　:QP和更新SPL权的时候+1
- FADMM ：不放松约束，但过滤掉权较小的样本（绿色）


dataset=thyroid kernel=gaussian
dataset=ionosphere kernel=linear
dataset=krvskp kernel=gaussian
dataset=image kernel=gaussian
dataset=clean1 kernel=gaussian
dataset=australian kernel=linear
dataset=phoneme kernel=gaussian
Accuracy
Number of queried instances
BMDR
random
ASPL
SPAL_add1
SPAL_fadmm

原目标:

$$\min_{f,w,v}\sum_{i=1}^{n_l}\left[v_i(y_i-f(x_i))^2+\lambda(\frac{1}{2}v_i^2-v_i)\right]+\sum_{j=1}^{n_u}\left[v_j\cdot w_j\left(\hat{y}_j-f\left(x_j\right)\right)^2+\lambda(\frac{1}{2}v_j^2-v_j)\right]+\mu(w^TK_1w+kw)+\gamma||f||^2$$

$$w_j\in[0,1],v_i\in[0,1]\quad\forall i=1\cdots n$$

参数w对应信息量和代表性，参数v对应样本的难易程度

新目标:

$$\min_{f,w}\sum_{i=1}^{n_l}[(y_i-f(x_i))^2]+\sum_{j=1}^{n_u}\left[w_j\left(\hat{y}_j-f\left(x_j\right)\right)^2+w_jT(f,\lambda)\right]+\mu(w^TK_1w+kw)+\gamma||f||^2$$

$$w_j\in[0,1]\quad\forall i=1\cdots n$$

仅有参数w，一项对应信息量，一项对应代表性，一项$T(f,\lambda)$对应样本的难易程度

思路：样本损失$\left(\hat{y}_j-f\left(x_j\right)\right)^2$超过$\lambda$时，$T(f,\lambda)$会快速上升，小于$\lambda$时，值较小，如sigmoid函数，指数函数，$xe^x$，hinge_loss等等

取指数函数:

$$\min_{f,w} \sum_{i=1}^{n_l} [(y_i - f(x_i))^2] + \sum_{j=1}^{n_u} \left[w_j \left(\hat{y}_j - f(x_j)\right)^2 + w_j \tau^{\left(\hat{y}_j - f(x_j)\right)^2 - \lambda} \right] + \mu(w^T K_1 w + kw) + \gamma ||f||^2$$

$$w_j \in [0,1] \quad \forall i = 1 \cdots n$$

Where τ is a constant

优化f时:

$$\min_{\theta} \sum_{i=1}^{n_l} \left(y_i - \sum_{x_k \in L} \theta_k k(x_k, x_i) \right)^2 + \sum_{j=1}^{n_u} \left[w_j \left(\sum_{x_k \in L} \theta_k k(x_k, x_j) \right)^2 + 2 w_j \left| \sum_{x_k \in L} \theta_k k(x_k, x_j) \right| \right.$$

取类似于Hinge_loss的分段函数:

$$\min_{f,w}\sum_{i=1}^{n_l}[(y_i-f(x_i))^2]+\sum_{j=1}^{n_u}\left[w_j\left(\hat{y}_j-f(x_j)\right)^2+w_j\max(0,\left(\hat{y}_j-f(x_j)\right)^2-\lambda)\right]+\mu(w^TK_1w+kw)+\gamma||f||^2$$

$$s.t\qquad w_j\in[0,1]\ \ \forall i=1\cdots n$$

$$\min_{f,w}\sum_{i=1}^{n_l}[(y_i-f(x_i))^2]+\sum_{j=1}^{n_u}\left[w_j\left(\hat{y}_j-f(x_j)\right)^2+w_j\xi_j\right]+\mu(w^TK_1w+kw)+\gamma||f||^2$$

$$w_j\in[0,1]\ \ \forall i=1\cdots n$$

$$\left(\hat{y}_j-f(x_j)\right)^2-\lambda)+\xi_j\geq 0$$

优化f时:

$$\min_{\theta,\xi}\sum_{i=1}^{n_l}\left(y_i-\sum_{x_k\in L}\theta_k k(x_k,x_i)\right)^2+\sum_{j=1}^{n_u}\left[w_j\left(\sum_{x_k\in L}\theta_k k(x_k,x_j)\right)^2+2w_j\left|\sum_{x_k\in L}\theta_k k(x_k,x_j)\right|+w_j\xi_j\right]+\gamma\theta^TK_{LL}\theta$$

$$f(x_j)^2+2|f(x_j)|+1-\lambda)+\xi_j\geq 0$$