Algorithm 1 DG (Duality Gap)

Input:

$X \in \mathbb{R}^{n\times p}$ ▷ The design matrix

$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors

$\beta \in \mathbb{R}^{n}$ ▷ Current β

$\lambda \in \mathbb{R}$ ▷ Grid element

1: $\epsilon \leftarrow X\beta - Y$

2: $f_\beta \leftarrow ||\epsilon||_2^2 + \lambda||\beta||_1$ ▷ Primal Objective function

3: $\alpha \leftarrow \frac{\lambda}{||2X^T\epsilon||_\infty}$

4: $\alpha_0 \leftarrow Y^T\epsilon$

5: $\mathcal{S} \leftarrow \min\{\max\{\alpha, \alpha_0\}, -\alpha\}$ ▷ Dual Point

6: $\widetilde{\nu} \leftarrow \frac{-2\mathcal{S}}{\lambda}\epsilon + \frac{2}{\lambda}Y$

7: $d_\nu \leftarrow \frac{1}{4}\lambda^2||\widetilde{\nu}||_2^2 - ||Y||_2^2$ ▷ Dual Objective function

return $f_\beta + d_\nu$

Algorithm 2 DGT (Duality Gap Target)

Input:

$\gamma \in \mathbb{R}$ ▷ ...

$C \in \mathbb{R}$ ▷ ...

$r_{StatsIt} \in \mathbb{N}$ ▷ Index of the current grid element (outer loop iteration number)

$n \in \mathbb{N}$ ▷ Number of rows in the design matrix $X \in \mathbb{R}^{n\times p}$

1: dgt $\leftarrow \gamma C^2 \frac{r_{StatsIt}^2}{n}$

return dgt

Algorithm 3 f_β

Input:

$X \in \mathbb{R}^{n\times p}$ ▷ The design matrix

$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors

$\beta \in \mathbb{R}^{n}$ ▷ Current β

1: $f \leftarrow X\beta - Y.$

return $||f||_2^2$

Algorithm 4 $f_{\widetilde{\beta}}$

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix

$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors

$\beta \in \mathbb{R}^{n}$ ▷ The k'th β vector

$\beta' \in \mathbb{R}^{n}$ ▷ The $k-1$'th β vector

$L \in \mathbb{R}$ ▷ The current Lipschitz constant, as computed by backtracking line search

$f \leftarrow X\beta - Y$.

$t_0 \leftarrow ||f||_2^2$

$\nabla f \leftarrow 2X^T f$

$\Delta_\beta \leftarrow \beta - \beta'$

$t_1 \leftarrow \nabla f^T \Delta_\beta$

$t_2 \leftarrow \frac{L}{2}||\Delta_\beta||_2^2$

return $t_0 + t_1 + t_2$

Algorithm 5 τ (Matrix-wise Soft-Thresholding / Proximal Operator)

Input:

$X \in \mathbb{R}^{n \times m}$ ▷ An arbitrary matrix

$\lambda \in \mathbb{R}$ ▷ The thresholding parameter

$\widetilde{X} \leftarrow X$ ▷ Make a copy of X.

for $\widetilde{x}_{i,j} \in \widetilde{X}$ **do**

  $\widetilde{x}_{i,j} \leftarrow \text{sign}(\widetilde{x}_{i,j})\,(|\widetilde{x}_{i,j}| - \lambda)^+$

end for

return $\widetilde{X}$

Algorithm 6 ISTA with backtracking line search and duality gap convergence criteria

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix
$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors
$\beta \in \mathbb{R}^{n}$ ▷ Starting vector
$L_0 \in \mathbb{R}$ ▷ Initial Lipschitz constant, used by backtracking line search
$\lambda \in \mathbb{R}$ ▷ Grid element
$\eta \in \mathbb{R}$ ▷ Step size when updating Lipschitz constant
$\mathcal{D} \in \mathbb{R}$ ▷ Duality gap target

```
β̃ ← β                                   ▷ Make a copy of β that will be updated during back tracking.
do
    β̃ ← τ(β − (1/L)∇f(X, Y, β̃, L))
    while f_β(X, Y, β̃) > f_β̃(X, Y, β̃, β, L) do
        L ← ηL
        β̃ ← τ(β − (1/L)∇f(X, Y, β))
    end while
    β ← τ(β − (1/L)∇f(X, Y, β, L))        ▷ Update β once L is sufficiently large.
while DG (X, Y, β, λ) > D

return β
```

$\widetilde{\beta} \leftarrow \beta$ ▷ Make a copy of β that will be updated during back tracking.
do
$\quad \widetilde{\beta} \leftarrow \tau(\beta - \frac{1}{L}\nabla f(X, Y, \widetilde{\beta}, L))$
$\quad$ **while** $f_{\beta}(X, Y, \widetilde{\beta}) > f_{\widetilde{\beta}}(X, Y, \widetilde{\beta}, \beta, L)$ **do**
$\quad\quad L \leftarrow \eta L$
$\quad\quad \widetilde{\beta} \leftarrow \tau(\beta - \frac{1}{L}\nabla f(X, Y, \beta))$
$\quad$ **end while**
$\quad \beta \leftarrow \tau(\beta - \frac{1}{L}\nabla f(X, Y, \beta, L))$ ▷ Update β once L is sufficiently large.
while DG $(X, Y, \beta, \lambda) > \mathcal{D}$

return β

Algorithm 7 Coordinate Descent with duality gap convergence criteria

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix
$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors
$\beta \in \mathbb{R}^{n}$ ▷ Starting vector
$\lambda \in \mathbb{R}$ ▷ Grid element
$\mathcal{D} \in \mathbb{R}$ ▷ Duality gap target

$\widetilde{\beta} \leftarrow \beta$ ▷ Make a copy of β
do
$\quad$ **for** $i \in 1, 2, \ldots, p$ **do**
$\quad\quad t \leftarrow \frac{\lambda}{2||X_i||_2^2}$ ▷ Scale grid element by norm of the i'th column of design matrix
$\quad\quad X_{-i} \leftarrow X_{\forall j \neq i}$ ▷ Take all columns of design matrix not equal to i
$\quad\quad \widetilde{\beta}_{-i} \leftarrow \widetilde{\beta}_{\forall j \neq i}$ ▷ Take all elements of predictors vectors not equal to i
$\quad\quad r \leftarrow \frac{X_i^T\left(Y - X_{-i}\widetilde{\beta}_{-i}\right)}{||X_i||_2^2}$ ▷ Compute the scaled residual
$\quad\quad \widetilde{\beta}_i \leftarrow \tau\left(r, t\right)$ ▷ Update the i'th element of Beta
$\quad$ **end for**
while DG $(X, Y, \widetilde{\beta}, \lambda) > \mathcal{D}$

return $\widetilde{\beta}$

Algorithm 8 Coordinate Descent with Lazy Evaluation

Input:
$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix
$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors
$\beta \in \mathbb{R}^{n}$ ▷ Starting vector
$\lambda \in \mathbb{R}$ ▷ Grid element
$\mathcal{D} \in \mathbb{R}$ ▷ Duality gap target
$\widetilde{\beta} \leftarrow \beta$ ▷ Make a copy of β
$R \leftarrow Y - X\widetilde{\beta}$ ▷ Initialize Intermediary Residual
do
  for $i \in 1, 2, \ldots, p$ **do**
    $t \leftarrow \frac{\lambda}{2||X_i||_2^2}$ ▷ Scale grid element by norm of the i'th column of design matrix
    if $\widetilde{\beta}_i \neq 0$ **then** $R \leftarrow R + X_i \widetilde{B}_i$
    end if
    $\widetilde{\beta}_i \leftarrow \tau\left(\frac{X_i^T R}{||X_i||_2^2}, t\right)$ ▷ Update the i'th element of Beta
    if $\widetilde{\beta}_i \neq 0$ **then** $R \leftarrow R - X_i \widetilde{B}_i$
    end if
  end for
while DG $(X, Y, \widetilde{\beta}, \lambda) > \mathcal{D}$

return $\widetilde{\beta}$

Algorithm 9 Coordinate Descent with standardized data

Input:
$X \in \mathbb{R}^{n \times p}$ ▷ The standardized $\bar{X}_i = 0, \sigma_{X_i} = 1$ design matrix
$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors
$\beta \in \mathbb{R}^{n}$ ▷ Starting vector
$\lambda \in \mathbb{R}$ ▷ Grid element
$\mathcal{D} \in \mathbb{R}$ ▷ Duality gap target
$\widetilde{\beta} \leftarrow \beta$ ▷ Make a copy of β
do
  for $i \in 1, 2, \ldots, p$ **do**
    $t \leftarrow \frac{\lambda}{2n}$ ▷ Scale grid element by norm of the i'th column of design matrix
    $X_{-i} \leftarrow X_{\forall j \neq i}$ ▷ Take all columns of design matrix not equal to i
    $\widetilde{\beta}_{-i} \leftarrow \widetilde{\beta}_{\forall j \neq i}$ ▷ Take all elements of predictors vectors not equal to i
    $r \leftarrow \frac{X_i^T\left(Y - X_{-i}\widetilde{\beta}_{-i}\right)}{n}$ ▷ Compute the scaled residual
    $\widetilde{\beta}_i \leftarrow \tau\left(r, t\right)$ ▷ Update the i'th element of Beta
  end for
while DG $(X, Y, \widetilde{\beta}, \lambda) > \mathcal{D}$

return $\widetilde{\beta}$

The next algorthim is a modified version of Coordinate Descent that seeks to reduce redunant computations as much as possible. This algorthim relies on the fact that many computation required by Coordinate Descent can be broken up into constant and non-constant parts. The constant parts of the computation can be performed ahead of time and stored for later use.

Of particular note is the scaled residual computation, which when written down naively reads:

$$r \leftarrow \frac{X_i^T\left(Y - X_{-i}\widetilde{\beta}_{-i}\right)}{||X_i||_2^2}.$$

Which we can re-write as,

$$r \leftarrow \frac{X_i^T Y}{||X_i||_2^2} - \frac{X_i^T X_{-i}}{||X_i||_2^2}\widetilde{\beta}_{-i}.$$

Note that since the design matrix X and the vector of predictors Y are fixed, the terms $\frac{X_i^T Y}{||X_i||_2^2}$ and $\frac{X_i^T X_{-i}}{||X_i||_2^2}$ do not change as the values of $\widetilde{\beta}$ are updated. Our strategy will be to compute these values for each column of X and store them in an array of size p, from which they will be access as $\widetilde{B}$ is updated.

Note for this algorthim we establish the convention that array of a given data type will be declared as follows:

$$(\ \textit{type}\ )[\ \#\ \textit{elements in array}\ ]$$

As an example an array of real numbers numbers of size $n \in \mathbb{N}$ would be written as:

$$(\ \mathbb{R}\ )[\ n\ ]$$

Algorithm 10 Coordinate Descent with Minimal Data Copying

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix
$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors
$\beta \in \mathbb{R}^{n}$ ▷ Starting vector
$\lambda \in \mathbb{R}$ ▷ Grid element
$\mathcal{D} \in \mathbb{R}$ ▷ Duality gap target

$\gimel \leftarrow (\mathbb{R})[p]$ ▷ Initialize array of size p to hold threshold parameters
$p_1 \in (\mathbb{R})[p]$ ▷ Blank array for part of residual computation
$p_2 \in (\mathbb{R}^{1 \times p})[p]$ ▷ Blank array to row vectors of size p which will be used for part of residual computation
for $i \in 1, 2, \dots, p$ **do**
  $\beth \leftarrow \frac{1}{||X_i||_2^2}$
  $\gimel[i] \leftarrow \frac{1}{2}\beth$
  $p_1[i] \leftarrow \beth X_i^T Y$
  $p_2[i] \leftarrow \beth X_i^T X$
end for
$\widetilde{\beta} \leftarrow \beta$ ▷ Make a copy of β
do
  for $i \in 1, 2, \dots, p$ **do**
    $\widetilde{\gamma} \leftarrow \widetilde{\beta}_j$ **if** $j \neq i$ **else** 0 ▷ Copy all of beta expect i'th element which is assigned to 0
    $r \leftarrow p_1[i] - p_2[i]\widetilde{\gamma}$ ▷ Compute the scaled residual
    $t \leftarrow \lambda\gimel[i]$ ▷ Compute threshold parameter
    $\widetilde{\beta}_i \leftarrow \tau\,(r, t)$ ▷ Update the i'th element of Beta
  end for
while DG $(X, Y, \widetilde{\beta}, \lambda) > \mathcal{D}$

return $\widetilde{\beta}$

Algorithm 11 FISTA with backtracking line search and duality gap convergence criteria

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix
$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors
$\beta \in \mathbb{R}^{n}$ ▷ Starting vector
$L_0 \in \mathbb{R}$ ▷ Initial Lipschitz constant, used by backtracking line search
$\lambda \in \mathbb{R}$ ▷ Grid element
$\eta \in \mathbb{R}$ ▷ Step size when updating Lipschitz constant
$\mathcal{D} \in \mathbb{R}$ ▷ Duality gap target
$y_{k-1} \in \mathbb{R}^{b}$ ▷ Beta vector from previous iteration of FISTA
$x_{k-1} \in \mathbb{R}^{b}$ ▷ Intermediary vector from previous iteration of FISTA

$y_k \leftarrow \beta$
do
    $y_{k-1} \leftarrow x_k$
    $t_k \leftarrow 0$
    $\widetilde{y_k} \leftarrow \tau(\beta - \frac{1}{L}\nabla f(X, Y, y_k))$
    while $f_\beta(X, Y, \widetilde{y_k}) > f_{\widetilde{\beta}}(X, Y, \widetilde{y_k}, y_k, L)$ **do**
        $L \leftarrow \eta L$
        $\widetilde{y_k} \leftarrow \tau(\beta - \frac{1}{L}\nabla f(X, Y, \beta))$
    end while
    $x_{k-1} \leftarrow x_k$
    $x_k \leftarrow \tau(\beta - \frac{1}{L}\nabla f(X, Y, \widetilde{y_k}))$
    $t_{k+1} = \frac{\left(1+\sqrt{1+4t_k^2}\right)}{2}$
    $y_k \leftarrow x_k + \frac{(t_k - 1)}{t_{k+1}}(x_k - x_{k-1})$
while DG $(X, Y, \beta, \lambda) > \mathcal{D}$

return $y_k, y_{k-1}, x_{k-1}, t_{k+1}$

Algorithm 12 λGRID

Input:

$X \in \mathbb{R}^{n \times m}$ ▷ The design matrix
$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors
$M \in \mathbb{N}$ ▷ The number of grid elements required

$r_{max} \leftarrow 2||X^T Y||_\infty$
$r_{min} \leftarrow \frac{1}{1000} r_{max}$
$\Delta_r \leftarrow (r_{max} - r_{min})$
Let $\Lambda \in \mathbb{R}^M$ ▷ Initialize empty array of size M
for $i \in [1, 2, \ldots, M]$ **do**
    $\delta_l \leftarrow \Delta_r \frac{i}{M-1} + r_{min}$ ▷ Compute linear step
    $\Lambda[i] \leftarrow 10^{\delta_l}$ ▷ Convert to logarithmic step
end for
return Λ

Algorithm 13 SCC: Stats Continutation Condition

Input:
$C \in \mathbb{R}$ ▷ The design matrix
$statsIt \in \mathbb{N}$ ▷ The vector of predictors
$\lambda \in \mathbb{R}$ ▷ Current grid element
$\Lambda \in \mathbb{R}^M$ ▷ Vector of grid elements
$X \in \mathbb{R}^{n \times p}$ ▷ Vector of grid elements
$\beta_s \in \mathbb{R}^{n \times M}$ ▷ Betas matrix
condition $\leftarrow$ false
for $i \in [1, 2, \ldots, statsIt]$ **do**
  $r_k \leftarrow \Lambda_k$
  $\Delta_\beta \leftarrow \beta_{statsIt} - \beta_i$
  check $\leftarrow \frac{n||\Delta_\beta||_\infty}{r_{statsIt}+r_k}$
  condition $\leftarrow$ condition & (check $\leq C$)
end for
return condition

Algorithm 14 FOS

Input:
$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix
$Y \in \mathbb{R}^n$ ▷ The vector of predictors
$\beta \in \mathbb{R}^n$ ▷ Starting vector
$L_0 \in \mathbb{R}$ ▷ Initial Lipschitz constant, used by backtracking line search
$M \in \mathbb{M}$ ▷ Number of grid elements
$\eta \in \mathbb{R}$ ▷ Step size when updating Lipschitz constant
$C \in \mathbb{R}$
$\gamma \in \mathbb{R}$
$\widetilde{X} \leftarrow \frac{1}{\sigma_X}(X - \mu_X)$. ▷ Normalize X to mean 0 and standard deviation 1.
$\widetilde{Y} \leftarrow \frac{1}{\sigma_Y}(Y - \mu_Y)$. ▷ Normalize Y.
$\Lambda \leftarrow \lambda\mathrm{GRID}(X, Y, M)$ ▷ Initialize grid elements
$\beta_s \in \mathbb{R}^{n \times m} = 0_{n,m}$. ▷ Initialize matrix of Betas to zero matrix
while statsCont & (statsCont < M) **do**
  $stats_{It} \leftarrow stats_{It} + 1$
  $\widetilde{\beta} \leftarrow \beta_{k-1}$ ▷ Initialize old beta vector with the k - 1'th Column of the Betas matrix.
  $r_{statsIt} \leftarrow \Lambda_k$ ▷ Extract the k'th grid element.
  if $DG(X, Y, \beta_k, r_{statsIt}) \leq DGT(\gamma, C, r_{statsIt}, n)$ **then**
    $\beta_k \leftarrow \beta_{k-1}$
  else
    $\beta_k \leftarrow \mathrm{ISTA}(X, Y, \beta_{k-1}, L_0, r_{statsIt}, \eta, gap)$
  end if
  $statsCont \leftarrow \mathrm{SCC}\ (C, statsIt, r_{statsIt}, \Lambda, X, \beta_s)$
end while
return $\beta_{statsIt-1}, \Lambda_{statsIt}, statsIt$

Algorithm 15 DP (Dual Point)

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix

$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors

$\beta \in \mathbb{R}^{p}$ ▷ Current β

$\lambda \in \mathbb{R}$ ▷ Grid element

1: $R \leftarrow Y - X\beta$

2: $\alpha \leftarrow \frac{1}{||X^T R||_\infty}$

3: $s \leftarrow \min\{\max\{\frac{Y^T R}{\lambda ||R||_2^2}, -\alpha\}, \alpha\}$

return sR

Algorithm 16 DG2 (Duality Gap for Problem 1)

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix

$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors

$\beta \in \mathbb{R}^{p}$ ▷ Current primal point

$\nu \in \mathbb{R}^{n}$ ▷ Current dual point

$\lambda \in \mathbb{R}$ ▷ Grid element

1: $f_\beta \leftarrow \frac{1}{2}||Y - X\beta||_2^2 + \lambda||\beta||_1$ ▷ Primal Objective function

2: $d_\nu \leftarrow \frac{1}{2}||Y||_2^2 - \frac{\lambda^2}{2}||\nu - \frac{Y}{\lambda}||_2^2$ ▷ Dual Objective function

return $f_\beta - d_\nu$

Algorithm 17 SAS (Safe Active Set)

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix

$c \in \mathbb{R}^{n}$ ▷ Center of the ball

$r \geq 0$ ▷ Radius of the ball

1: $\mathcal{A} \leftarrow \emptyset$ ▷ Initialize Active Set With Empty Set

2: **for** $j \in \{1, \ldots, p\}$ **do**

3:     **if** $|X_j^T c| + r\|X_j\|_2 \geq 1$ **then**

4:         $\mathcal{A} \leftarrow \mathcal{A} \cup \{j\}$

5:     **end if**

6: **end for**

return $\mathcal{A}$

Algorithm 18 CDSR (Coordinate Descent With Lazy Evaluation and Screening Rule)

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix
$Y \in \mathbb{R}^{n}$ ▷ The vector of predictors
$\beta \in \mathbb{R}^{p}$ ▷ Starting vector
$\lambda \in \mathbb{R}$ ▷ Grid element
$\mathcal{D} \in \mathbb{R}$ ▷ Duality gap target

$\widetilde{\beta} \leftarrow \beta$ ▷ Make a copy of β
$R \leftarrow Y - X\widetilde{\beta}$ ▷ Initialize Intermediary Residual
$\mathcal{A} \leftarrow \{1, \dots, p\}$ ▷ Initialize Active Set
$optimCont \leftarrow true$
while optimCont **do**
  $\nu \leftarrow \mathrm{DP}(X_{\mathcal{A}}, Y, \widetilde{\beta}_{\mathcal{A}}, \lambda)$ ▷ Dual point
  $\mathcal{G} \leftarrow \mathrm{DG2}(X_{\mathcal{A}}, Y, \widetilde{\beta}_{\mathcal{A}}, \nu, \lambda)$ ▷ Duality gap
  $\mathcal{A} \leftarrow \mathrm{SAS}(X, \nu, \sqrt{\frac{2}{\lambda^2}\mathcal{G}})$ ▷ Safe Active Set
  if $\mathcal{G} \leq \mathcal{D}$ **then**
    $optimCont \leftarrow false$
  else
    for $i \in \mathcal{A}$ **do**
      $t \leftarrow \frac{\lambda}{||X_i||_2^2}$ ▷ Scale grid element by norm of the i'th column of design matrix
      if $\widetilde{\beta}_i \neq 0$ **then** $R \leftarrow R + X_i\widetilde{\beta}_i$
      end if
      $\widetilde{\beta}_i \leftarrow \tau\left(\frac{X_i^T R}{||X_i||_2^2}, t\right)$ ▷ Update the i'th element of Beta
      if $\widetilde{\beta}_i \neq 0$ **then** $R \leftarrow R - X_i\widetilde{\beta}_i$
      end if
    end for
  end if
  $\widetilde{\beta}_{\mathcal{A}^c} = 0$ ▷ Set to 0 coefficients not in $\mathcal{A}$
end while
return $\widetilde{\beta}$

Note that Algorithm 18 solves the problem

$$\arg\min_{\beta \in \mathbb{R}^p} \frac{1}{2}||Y - X\beta||_2^2 + \lambda||\beta||_1 \tag{1}$$

Algorithm 19 FOS With Screening Rule

Input:

$X \in \mathbb{R}^{n \times p}$ ▷ The design matrix
$Y \in \mathbb{R}^n$ ▷ The vector of predictors
$M \in \mathbb{N}$ ▷ Number of grid elements
$C > 0$
$\gamma > 0$

$\widetilde{X} \leftarrow \frac{1}{\sigma_X}(X - \mu_X)$ ▷ Normalize X to mean 0 and standard deviation 1.
$\widetilde{Y} \leftarrow \frac{1}{\sigma_Y}(Y - \mu_Y)$ ▷ Normalize Y.
$\Lambda \leftarrow \lambda\text{GRID}(\widetilde{X}, \widetilde{Y}, M)$ ▷ Initialize grid elements
$\beta_s \in \mathbb{R}^{p \times M} \leftarrow 0_{p,M}$ ▷ Initialize matrix of Betas to zero matrix
$statsCont \leftarrow true$
$statsIt \leftarrow 1$
while statsCont & (statsIt $< M$) **do**
  $statsIt \leftarrow statsIt + 1$
  $\mathcal{D} \leftarrow DGT(\gamma, C, \Lambda_{statsIt}, n)$ ▷ Duality gap target
  $\beta_{statsIt} \leftarrow \text{CDSR}(\widetilde{X}, \widetilde{Y}, \beta_{statsIt-1}, \Lambda_{statsIt}/2, \mathcal{D}/2)$
  $statsCont \leftarrow \text{SCC}\ (C, statsIt, \Lambda_{statsIt}, \Lambda, X, \beta_s)$
end while
return $\beta_{statsIt-1}, \Lambda_{statsIt-1}, statsIt$