have complex aliasing, like the Plackett-Burman designs, they are best analyzed with a regression subset procedure as described above. Some authors such as Lin (1999) have suggested the use of forward stepwise regression to identify an appropriate model for data arising from a design with complex aliasing. To illustrate this proceedure in R, consider again the data for the cast fatigue experiment. The `step` function could be used as shown below to identify a model.

```
> null <- lm( y ~ 1, data = castfr )
> up <- lm( y ~ (.)^2, data = castfr )
> step( null, scope = list(lower = null, upper = up),
+ direction = "forward", steps=4)
```

This code operates on the reduced data frame `castfr` containing only the main effect columns and the response from the Plackett-Burman design created in the code shown earlier. `null` defines a model with the minimum number of terms to be considered. In this case that is a model with only an intercept. `up` defines the set of all terms to be considered for the model. In this statement the model formula, `y ~ (.)^2`, creates all main effects and two-factor interactions. In the call of the `step` function, `null` is specified as the starting model and the option `steps=4` specifies the maximum number of forward steps. This should normally be set to 1/3 of the number of runs in the design because due to the effect sparsity principle there will rarely be more than that number of important effects. The step function uses the AIC or equivalently Mallows's $C_p$ to decide whether additional terms should be added to the model. Running this code results in two forward steps. In the first step, main effect $F$ is added to the model, and in the second and final step, main effect $D$ is added to the model.

In the analysis of the cast fatigue experiment discussed earlier, two alternative two-variable models were plausible. Authors of the original article describing the experiment thought the model including main effects F and D was appropriate, but using all-subsets regression, the two-variable model ($F$ and $FG$) fit the data much better. This model does not include any interactions that do not involve at least one of the main effects in the model (effect heredity).

In some cases, there there may be more than one model that fits the data well. A forward selection procedure only identifies one of these models. An all-subsets selection procedure can identify all of the models that fit the data well. However, the concern when using the all-subsets regression is an increase in computations, especially for designs with a large number of factors.

To alleviate the concern about increased computations required for all-subsets regression, Hamada and Wu (1992) proposed a more involved iterative stagewise forward stepwise regression approach, guided by the principle of effect heredity, that overcomes some of the objections to using a straight forward regression. Jones and Nachtsheim (2011) proposed a simpler approach to forward regression that also forces effect heredity. Their approach, incorporated with the "Combine" option in the JMP forward stepwise regression, requires that any forward step that enters an interaction effect to the model also enters the main effects involved in that interaction. This avoids finding models that do not obey effect heredity (a situation that occurred less than 1% of the time in Li *et al.*'s (2006) study of 113 published factorial experiments).

The function `HierAFS()` in the package `daewr` performs a stepwise forward regression that also includes both main effects involved in any interaction added to the model in order to insure model hierarchy like the "Combine" option in the JMP forward regression. An example of the use of this function with the data from the cast fatigue experiment is shown below.

```
> des <- castfr[ ,c(1, 2, 3, 4, 5, 6,7 )]
> y <- castfr[ ,8]
> library(daewr)
> HierAFS(y,des,m=0,c=7,step=2)
      formula    R2
1   y~F+G+F:G 0.910
2 y~D+F+G+F:G 0.937
```

In the `HierAFS` function call the first argument, `y` is the vector of responses, `des` is the design matrix, `m=0` is the number of three level factors in the design, `c=7` is the number of two-level factors in the design, and `step=2` is the number of forward steps to make. This number is usually determined by the number of steps just before the increase in $R^2$ becomes negligible or just before the last term to enter the model is not significant.

In the first step of the forward regression, it can be seen that the FG interaction entered the model and the main effects F and G were automatically included. In the second step, the main effect D entered the model. The code and output below show the details of the first and second steps by printing the summary of the two models fit by the `lm` function.

```
> mod1<-lm(y~F+G+F:G,data=des)
> summary(mod1)


Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.73025    0.07260  78.930  7.4e-13 ***
F            0.45758    0.07260   6.303 0.000232 ***
G            0.09158    0.07260   1.261 0.242669
F:G         -0.45875    0.07260  -6.319 0.000228 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2515 on 8 degrees of freedom
```

```
Multiple R-squared:  0.9104,Adjusted R-squared:  0.8767
F-statistic: 27.08 on 3 and 8 DF,  p-value: 0.0001531
```

```
> mod2<-lm(y~D+F+G+F:G,data=des)
> summary(mod2)
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.73025    0.06516  87.939 6.48e-12 ***
D           -0.11831    0.06911  -1.712 0.130661
F            0.45758    0.06516   7.022 0.000207 ***
G            0.09158    0.06516   1.405 0.202676
F:G         -0.41931    0.06911  -6.067 0.000507 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2257 on 7 degrees of freedom
Multiple R-squared:  0.9368,Adjusted R-squared:  0.9007
F-statistic: 25.95 on 4 and 7 DF,  p-value: 0.0002713
```

Since the term, `D`, entered at the second step was not significant, the forward regression should have stopped after the first step, and the hierarchical model shown in the summary of `mod1` is the same model for the data found by the all-subsets regression.

In Plackett-Burman or alternative screening designs with a large number of factors, the computational effort required to do an all-subsets regression to search for the most appropriate model may be prohibitive. The `HierAFS` function often finds an appropriate model with much less computation.

Other methods have been proposed to limit the number of candidate terms for a model search. Lawson (2002) proposed limiting the interaction candidates for an all-subsets regression based on the alias structure of the design. He provides a SAS (Institute, 2012) macro to implement the method. Box and Meyer (1993) proposed a Bayesian approach for identifying an appropriate model, and Chipman *et al.* (1997) proposed using the Bayesian stochastic search algorithm that incorporates the effect heredity principle through heredity priors that capture the relation between the importance of an interaction term and the main effects from which it is formed. Woodward (2011) has incorporated Chipman *et al.*'s (1997) search algorithm using the public domain program WinBUGS (Spiegelhalter *et al.*, 1999) that can be called through his Excel add-in BugXLA. Wolters and Bingham (2011) proposed a simulated annealing model search with associated graphs that are used to identify good canditate models and assess the degree of uncertainty in model selection. They also provide MATLAB® (MATLAB, 2010) code to implement this strategy. For designs where the CPU time required to run the `regsubsets` function (as illustrated in the analysis of the cast fatigue experiment) is too long, one of these alternate analysis strategies should be employed.

In many situations, model robust screening designs with complex aliasing may reduce the total number of experiments required to identify the important main effects and two factor interactions. However, the data analysis with a regression subsetting procedure can be more involved than the simple analysis of $2^{k-p}$ designs. Also, if a three-factor interaction like that shown in Figure 6.10 is important, it would be very difficult to detect with regression subset selection. Therefore both traditional $2^{k-p}$ designs and model robust designs with complex aliasing should have their place in an experimenter's toolbox.

## 6.7 Mixed Level Factorials and Orthogonal Arrays (OAs)

In the preliminary stage of experimentation, where the objective may be to determine which factors are important from a long list of candidates, two-level fractional factorial designs or Plackett-Burman designs are often appropriate. If a factor has quantitative levels, the two levels are denoted symbolically by (−) and (+), where (−) represents the lowest level the experimenter would consider, and (+) represents the highest level the experimenter would consider. The high and low are usually spread out as far as feasibly possible in order to accentuate the signal or difference in response between the two levels. If a factor has qualitative levels, the (−) and (+) designations are arbitrary, but the two levels chosen normally would be two that the experimenter believes should result in the maximum difference in response.

Sometimes, however, two levels for each factor may not be adequate. In cases where the experimenter would like to consider nonlinear effects of quantitative factors or qualitative factors with more than two alternatives, two-level fractional designs will not be suitable. For example, Fannin *et al.* (1981) report an experiment investigating the effects of four three-level factors and two two-level factors upon the rate of bacterial degradation of phenol for the purpose of evaluating the fate of chemicals in aquatic ecosystems. A full factorial would require $3^4 \times 2^2 = 324$ experiments; however, the study was completed using only a fraction of these runs by utilizing a mixed level fractional factorial design based on an orthogonal array. Taguchi (1986) describes an experiment to determine the factors that affect the durability of an auto clutch spring. The factors and levels are shown in the table on the next page.

The levels of factors $A$, $C$, $F$, and $G$ represented discrete alternatives that were of interest. Factors $B$, $D$, and $E$ were continuous factors and three levels were included in order to determine whether there was a curvilinear relation between these factor levels and durability of the clutch springs. There was also interest in the interaction between factors $D$ and $F$ and the interaction